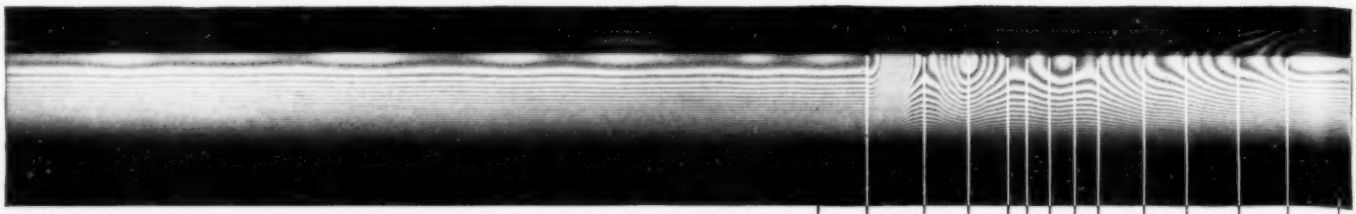
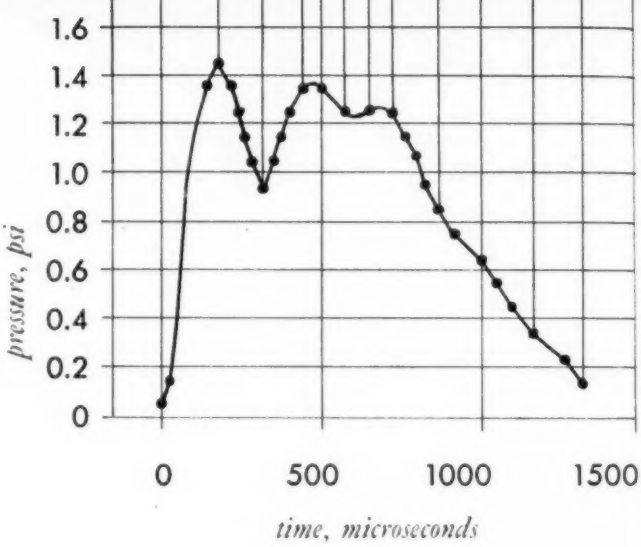
JOURNAL OF APPLIED PHYSICS

FAST WORK BY

NEWTON RINGS

capturing pressure transients





Light, the ideal yardstick for measuring rapid events, creates a pressure-transient measuring system of unusual capability in the Beckman & Whitley Model 216 Pressure Recorder. Having stable linear calibration, the unit can cover pressure ranges from 3 to 50,000 psi, with a 1 per cent full-scale accuracy in both pressure and time, and exhibits a frequency response from 10,000 to 50,000 cps.

To cover the widest possible variety of application requirements, the recorder is engineered so it can be used in three ways: (1) exposed to free air manifestations, (2) attached directly to a pressure vessel, or (3) fed by a hydraulic line from the pressure source.

The instrument is organized into two units, shown below. Transducer, optical system, and camera with timing-pip system are in the unit at left. The control unit, right, permits records to be made remotely.

Rapid response, accurate measurement of small displacements, high sensitivity, and stable zero reference points are all achieved by recording on moving film the motion of small arcs of optical-interference rings responding to the influence of the pressure fluctuations under measurement. Such fringe changes can be seen in the typical record above, while underneath, the simple procedure of translating the pattern into a plot of pressure-versus-time is outlined.

SEND FOR FURTHER DETAILS.

Beckman & Whitley INC.

instruments for physical research

SAN CARLOS 8, CALIFORNIA

Visit us at Booth #C-515-1, 10th Annual Instrument-Automation Conference and Exhibit, Los Angeles, September 12 to 16.

JOURNAL OF APPLIED PHYSICS

Vol. 26, No. 8 AUGUST, 1955

In This Issue

Page

Slowing Down of Neutrons by Hydrogenous Moderators . . . . . H. Hurwitz, Jr., and P. F. Zweifel 923

Tracer Diffusion of Iron in Stainless Steel . . . . . V. Linnenbom, M. Tetenbaum, and C. Cheek 932

Measurement of Excess Ba in Practical Oxide Coated Cathodes . . . . . L. A. Wooten, George E. Moore, and W. G. Guldner 937

Excess Ba Content of Practical Oxide Coated Cathodes and Thermionic Emission . . . . . George E. Moore, L. A. Wooten, and J. Morrison 943

Forward Characteristic of Germanium Point Contact Rectifiers . . . . . Melvin Cutler 949

Similarity Solution for a Spherical Shock Wave . . . . . Richard Latter 954

Geometrical Acoustics. II. Diffraction, Reflection, and Refraction of a Weak Spherical or Cylindrical Shock at a Plane Interface . . . . . K. O. Friedrichs and Joseph B. Keller 961

X-Ray Path through Calcite in Laue Diffraction . . . . . George L. Rogosa and Guenter Schwarz 967

Weak Detonations and Condensation Shocks . . . . . W. H. Heybey and S. G. Reed, Jr. 969

Preparation of Thin Magnetic Films and Their Properties . . . . . M. S. Blois, Jr. 975

On the Reflection of Shock Waves from an Open End of a Duct . . . . . George Rudinger 981

Domain Configurations and Crystallographic Orientation in Grain-Oriented Silicon Steel . . . . . W. S Paxton and T. G. Nilan 994

Maximum Superheating of Water as a Measure of Negative Pressure . . . . . Lyman J. Briggs 1001

Photoelectric Work Functions of the Borides of Lanthanum, Praseodymium, and Neodymium . . . . . Richard W. Decker and Dean W. Stebbins 1004

Influence of Space Charge on the Potential Distribution in Mass Spectrometer Ion Sources . . . . . W. M. Brubaker 1007

Slowing Down Distribution of U^{235} Fission Neutrons from a Point Source in Light Water . . . . . J. E. Hill, L. D. Roberts, and T. E. Fitch 1013

Slowing Down Distribution to Indium Resonance of U^{235} Fission Neutrons from a Point Fission Source in Two Aluminum Light Water Mixtures . . . . . L. D. Roberts, J. E. Hill, and T. E. Fitch 1018

Rectification Properties of Metal Semiconductor Contacts . . . . . E. H. Borneman, R. F. Schwarz, and J. J. Stickler 1021

Plasma Oscillations in Electron Beams . . . . . K. G. Hernqvist 1029

Power Flow in Electron Beams . . . . . L. R. Walker 1031

On the Space-Charge-Limited Current between Nonsymmetrical Electrodes . . . . George D. O'Neill 1034

Correlation of Diffraction and Transmission Experiments for X-Ray and Neutron Elastic Scattering . . . . . D. T. Keating and J. J. Antal 1041

Radiation Induced Changes in the Electrical Resistivity of α Brass . . . . . D. B. Rosenblatt, R. Smoluchowski, and G. J. Dienes 1044

Dependence of the Coercive Force on the Density of Some Iron Oxide Powders . . . . . A. H. Morrish and S. P. Yu 1049

Letters to the Editor:

Note on Electron Diffraction Patterns of CuO . . . . R. D. Heidenreich and K. H. Storks 1056

Sputtering of Metal Single Crystals by Ion Bombardment . . . . . Gottfried K. Wehner 1056

Interpretation of the Strain Broadening Components in X-Ray Diffraction Patterns . . . . . J. C. Bowman, J. A. Krumhansl, and J. R. Stock 1057

Thermal Effects in Point-Contact Rectifiers . . . . . R. E. Burgess 1058

Books Reviewed 1058

Books Received 1060

Announcements 1060

Index to Advertisers xxii

Visit us at Booth #C-315-1, 10th Annual Instrument-Automation Conference and Exhibit, Los Angeles, September

JOURNAL OF APPLIED PHYSICS

A journal of applications of physics to other sciences and to industry

Published by the

American Institute of Physics

GOVERNING BOARD

Allen V. Astin
Robert F. Bacher
H. A. Bethe
J. W. Buchta
S. A. Goudsmit
Deane B. Judd
Hugh S. Knowles
W. H. Markwood, Jr.
William F. Meggers
Philip M. Morse
Brian O'Brien
Harry F. Olson
R. F. Paton
Eric Rodgers
Ralph A. Sawyer
Frederick Seitz
William Shockley
H. D. Smyth
J. H. Van Vleck
Mark W. Zemansky

OFFICERS AND STAFF

Frederick Seitz
Chairman, Governing Board

Henry A. Barton
Director

George B. Pegram
Treasurer

Wallace Waterfall
Executive Secretary

Theodore Vorburger
Advertising Manager

Ruth F. Bryans
Publication Manager

Edith I. Neftel
Circulation Manager

Kathryn Setze
Assistant Treasurer

Melvin Loos
Consultant on Printing

Editor
Robert L. Sproull

Editorial Board

Maurice A. Biot
Walker Bleakney
Wallace J. Eckert
Elmer Hutchisson
D. O. North
Louis A. Pipes
W. R. Smythe
Arthur V. Tobolsky
A. M. Weinberg

THE JOURNAL OF APPLIED PHYSICS is devoted to the application of physics to other sciences, to engineering, and to industry. Previous to 1937, this journal was called PHYSICS.

The JOURNAL OF APPLIED PHYSICS welcomes articles of two distinct kinds: (1) Descriptions of new results in experimental or theoretical physics which are of general interest to applied physicists. The author should explain his contribution in such a way that a physicist or engineer in another field of specialization can appreciate its objectives and essential features. (2) Critical reviews of particular areas of applied physics. Authors contemplating writing such reviews should correspond with the Editor before they begin writing.

The JOURNAL OF APPLIED PHYSICS welcomes letters to the editor which describe in a few words the content of laboratory reports and theses and which tell readers how they may obtain copies of such reports.

Manuscripts should be submitted to the Editor, Rockefeller Hall, Cornell University, Ithaca, New York. Authors should read the "Information for Contributors," facing page 1 of the January issue, before writing articles or letters. The *Style Manual* of the American Institute of Physics is also helpful to writers preparing manuscripts for journals published by the American Institute of Physics. It can be obtained for $1.00 from the American Institute of Physics, 57 East 55th Street, New York 22, New York.

Corrected proofs and all correspondence concerning papers in the process of publication should be addressed to the Publication Manager, American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	14.00

Back Number Prices

Yearly back number rate when complete year is available: $15.00.
Single copies: $1.50 each.

Changes of address should be sent to the American Institute of Physics.

Advertising rates supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Journal
of
Applied Physics

Volume 26, Number 8 August, 1955

Slowing Down of Neutrons by Hydrogenous Moderators*

H. Hurwitz, Jr., and P. F. Zweifel
Knolls Atomic Power Laboratory,† Schenectady, New York
(Received February 2, 1955)

Methods of treating the Boltzmann equation for the case of a sinusoidally distributed source of fast neutrons in a material containing hydrogen are discusssed and compared. The effects of the simplifying assumptions commonly employed in reactor calculations are analyzed. A technique which is particularly suitable for accurate numerical integration of the Boltzmann equation in the above case is described. Numerical results are being obtained by this method, and will be presented and compared with results obtained from other approximate methods in a subsequent paper.

I. INTRODUCTION

THE extensive use of water as a moderator and coolant in small, highly enriched reactors has created great interest in the development of methods for the theoretical treatment of the neutron slowing-down problem in mixtures of hydrogen with heavier elements. Although the reactors may be heterogeneous, in the sense that the water and the fuel elements are mechanically separated, the inhomogeneities often have dimensions which are small compared to a neutron mean free path. This means that from the neutron transport standpoint, the assemblies may be considered as homogeneous to a first approximation, with the effects of the inhomogeneities being treated separately as perturbations.

For reactors not containing hydrogen, the Fermi age approximation[1] has been used extensively in solving the neutron transport problem.[2] Because of the large energy change suffered by a neutron in collision with a proton, however, this method is not directly applicable to hydrogen-moderated systems. In this paper several different methods for the approximate solution of the Boltzmann equation in the presence of hydrogen are discussed and compared. In a subsequent paper the results of numerical calculations of the slowing-down problem in hydrogenous media will be described.

In order to facilitate comparisons of the various methods, we consider the slowing-down problem in an infinite region in which the source energy spectrum is independent of position and the source strength varies spatially as e^{ikx}. These assumptions allow the spatial dependence to be factored out of the Boltzmann equation so that only the variation of flux with energy and angle need be considered. Such a problem is of more than academic interest for two reasons: first, a source having this spatial dependence may be considered as corresponding to a single Fourier component of an arbitrary source; and secondly, in reactors the spatial distribution of fissions (and thus the source) is often approximately sinusoidal. In this latter case, the wave number, k, of the source distribution is equal to B where B^2 is the buckling of the reactor.[3]

The methods considered here all involve the expansion of the neutron flux and the scattering function[4,5] in series of Legendre polynomials. The methods differ, however, in the approximation employed to obtain a finite set of coupled integral equations from the original

* A preliminary report of this work was given at the Washington, D. C. meeting of the American Physical Society in May, 1954; see Phys. Rev. **95**, 660 (1954).

† Operated for the U. S. Atomic Energy Commission by the General Electric Company.

[1] R. E. Marshak, Revs. Modern Phys. **19**, 185 (1947), or see reference 5.

[2] R. Ehrlich and H. Hurwitz, Jr., Nucleonics **12**, No. 2, 23 (1954).

[3] See, for example, S. Glasstone and M. C. Edlund, *The Elements of Nuclear Reactor Theory* (D. Van Nostrand Company, Inc., New York, 1952). Our calculations employ a source distribution e^{ikx} rather than $\sin kx$ since the equation is not factorable with the latter distribution, although all scalar quantities such as the integral of flux over angle and the slowing-down density have the same spatial distribution as the source in either case.

[4] Marshak, Brooks, and Hurwitz, Nucleonics **4**, No. 5, 10 (1949).

[5] Marshak, Brooks, and Hurwitz, Nucleonics **5**, No. 1, 53 (1949).

infinite set which is obtained when the Legendre polynomial series are put into the Boltzmann equation. In the standard spherical harmonic P_L approximation, the coefficients of the Legendre polynomials of degree greater than L in the expansion of the neutron flux are neglected. A somewhat more accurate procedure is to neglect the higher expansion coefficients of the scattering function rather than of the flux.[6,7] We refer to this as the "B_L approximation."

In addition to the approximation necessary to obtain a finite set of coupled integral equations, it is also frequently desirable to employ approximate methods in solving the integral equations. In particular, some of the energy integrals are treated by expanding the coefficients of the flux in a Taylor's series in the energy variable.[8] Terms arising from the heavier elements are, in fact, always treated in this manner.

The Fermi age-diffusion approximation, which is used when hydrogen is not present, is a modification of the P_1 approximation. In the energy integrals involving the isotropic component of the flux, N_0, only the first two terms in the Taylor series expansion of N_0 in the energy variable are retained, while in the integrals involving the P_1 component (N_1) only one term is used. The simplest improvement which can be made in the age-diffusion approximation when hydrogen is present is to avoid the Taylor expansion of N_0 in the integrals involving the hydrogen scattering cross section. If the other approximations in the age-diffusion equation are retained, this procedure has the effect of replacing the age-diffusion equation for N_0 by two coupled equations [see Eqs. (37) and (39)]. This procedure, as applied to reactor calculations, is known as the Selengut-Goertzel (S.G.) method.[9] It in effect makes use of the correct distribution of energy loss suffered by a neutron in a collision with a proton, but neglects the correlation between energy loss and angle of scattering.

The numerical calculations to be presented in a subsequent paper investigate separately the two types of errors in the S.G. and similar methods, namely the neglect of the higher components in the angular distribution and the approximate treatment of the integral equations in energy.

II. MATHEMATICAL ANALYSIS

1. General Remarks

The one-dimensional Boltzmann equation with a neutron source whose spatial dependence is of the form e^{ikx} may be written[8]

$$\sigma_T(u)(1-ik\mu/\sigma_T(u))\sum_{l=0}^{\infty}\frac{2l+1}{4\pi}N_l(k,u)P_l(\mu)$$

$$=\sum_\gamma\sum_{j'=0}^{\infty}\frac{2j'+1}{4\pi}J_{j'}{}^{\gamma}(k,u)P_{j'}(\mu)+\frac{1}{4\pi}S(u). \quad (1)$$

We adopt the notation of A, with two minor exceptions. First, the neutron flux is a function of the variable k rather than of x because of the assumed spatial dependence of the source; and second, the integrals $J_{j'}{}^{\gamma}$ are summed over all types of nuclei present, as here the presence of more than one moderating material is considered.

Reviewing briefly some of the discussion of A, $J_j{}^{\gamma}$ is defined by the relation

$$J_j{}^{\gamma}(k,u)=\int_{u-\Delta^\gamma}^{u}N_j(k,u')B_j{}^{o\gamma}(u',U)du',\ U=u-u', \quad (2)$$

where $\Delta^\gamma=\ln(1-\gamma/1+\gamma)^2$ is the maximum possible logarithmic energy (lethargy) change possible in an elastic collision between a neutron and a nucleus whose mass is m_n/γ, m_n being the mass of the neutron.

The $B_L{}^{o\gamma}$ are the Legendre components of the cross section[10] for a neutron of lethargy u' to scatter through an angle $\cos^{-1}\mu_o$ with a lethargy change U. (The index "o" indicates laboratory frame of reference.) Dropping the index γ,

$$B_j{}^{o}(u',U)=2\pi\int_{-1}^{1}d\mu_o P_j(\mu_o)\sigma^o(\mu_o,u',U). \quad (3)$$

Since μ_o and U are related by momentum considerations (A-11), one can write

$$B_j{}^{o}(u',U)=2\pi\sigma^o(u',\mu_o(U))P_j(\mu_0(U))\left(-\frac{d\mu_o(U)}{dU}\right), \quad (4a)$$

or

$$B_j{}^{o}(u',U)=2\pi\sigma^c(u',\mu_c(U))P_j(\mu_0(U))\left(-\frac{d\mu_c(U)}{dU}\right), \quad (4b)$$

where σ^o and σ^c are the differential scattering cross sections in the laboratory and center-of-mass systems respectively, and $\cos^{-1}\mu_c$ is the scattering angle in the latter system.

If $\sigma^c(u',\mu_c)$ is expanded in a series of Legendre polynomials,

$$\sigma^c(u',\mu_c)=\sum_{l'=0}^{\infty}\frac{2l'H}{4\pi}B_{l'}{}^{c}(u')P_{l'}(\mu_c), \quad (5)$$

one can write

$$B_j{}^{o}(u',U)=\sum_{l'=0}^{\infty}T_{jl'}(U)B_{l'}{}^{c}(u'), \quad (6)$$

[6] Bethe, Tonks, and Hurwitz, Phys. Rev. **80**, 11 (1950).

[7] D. Selengut, private communication. Waller and G. C. Wick (see reference 11 and 13 also used procedures similar to that used here.

[8] P. F. Zweifel and H. Hurwitz, Jr., J. Appl. Phys. **25**, 1241 (1954). We refer to this paper as "A", and to equations therein as, e.g., A-11.

[9] G. Goertzel (unpublished).

[10] All cross sections in this paper are assumed to be macroscopic, i.e., the cross section per nucleus multiplied by the number of nuclei per unit volume.

where the $T_{jl'}$ are the elements of a matrix defined in (A-25).

For isotropic scattering in the center-of-mass system,

$$B_{l'}{}^{c}(u')=\sigma_s(u')\delta_{l'0}, \tag{7}$$

so that one need consider only the $T_{j0}(U)$:

$$T_{j0}(U)=-\tfrac{1}{2}P_j(\mu_c(U))\frac{d\mu_c(U)}{dU}, \tag{8}$$

and

$$B_j{}^{o}(u',U)=\sigma_s(u')\frac{(1+\gamma)^2}{4\gamma}e^{-U}$$

$$\times P_j\left(\frac{1}{2\gamma}[(1+\gamma)e^{-U/2}-(1-\gamma)e^{U/2}]\right). \tag{9}$$

For hydrogen, $\gamma=1$, and Eq. (9) takes a particularly simple form:

$$B_j{}^{oH}(u',U)=\sigma_s{}^{H}(u')e^{-U}P_j(e^{-U/2}). \tag{10}$$

The use of Eq. (10) is permissible because $n-p$ scattering is isotropic at energies relevant to reactor calculations.

For heavier moderators ($\gamma<1$), the center-of-mass scattering is, in general, not isotropic. One then makes the Taylor's series expansion discussed in Sec. I (see A-30), obtaining

$$J_j{}^{\gamma}(k,u)=\sum_{l',n=0}^{\infty}T_{jl'}{}^{\gamma n}\frac{\partial^n}{\partial u^n}(N_j(k,u)B_{l'}{}^{\gamma c}(u)), \tag{11}$$

where

$$T_{jl'}{}^{\gamma n}=\frac{(-)^n}{n!}\int_0^{\Delta\gamma}U^nT_{jl'}{}^{\gamma}(U)dU. \tag{12}$$

Thus, one can write

$$\sum_\gamma J_j{}^{\gamma}(k,u)=\int_0^u N_j(k,u')\sigma_s{}^{H}(u')e^{-U}P_j(e^{-U/2})du'$$

$$+\sum_{\gamma\neq1}\sum_{l',n}T_{jl'}{}^{\gamma n}\frac{\partial^n}{\partial u^n}(N_j(k,u)B_{l'}{}^{\gamma c}(u)). \tag{13}$$

In practice it is usually sufficient to keep only those terms for $n=0$, 1 when $j=0$, and $n=0$ when $j>0$.

2. The P_L Approximation

The P_L approximation is obtained if Eq. (1) is multiplied by $P_j(\mu)$, and integrated over solid angle. An infinite set of coupled integral equations for the N_j is obtained:

$$\sigma_T(u)N_j(k,u)-\frac{ik}{2j+1}[(j+1)N_{j+1}(k,u)+jN_{j-1}(k,u)]$$

$$=\sum_\gamma J_j{}^{\gamma}(k,u)+S(u)\delta_{j0}, \tag{14}$$

where the source has been assumed isotropic. The assumption is then made that $N_{L+1}=0$, so that the first $L+1$ equations can be used to calculate $N_0, N_1, \cdots, N_L$.

3. The B_L Approximation

To obtain the B_L approximation, one divides Eq. (1) by $(1-ik\mu/\sigma_T)$ before multiplying by $P_j(\mu)$ and performing the angular integration. A different set of simultaneous integral equations for the N_j is obtained, namely,

$$\sigma_T(u)N_j(k,u)=\sum_{l=0}^{\infty}(2l+1)A_{lj}(k,u)$$

$$\times(\textstyle\sum_\gamma J_j{}^{\gamma}(k,u)+S(u)\delta_{j0}). \tag{15}$$

Here

$$A_{lj}(k,u)=A_{jl}(k,u)=\frac{1}{2}\int_{-1}^{1}d\mu\frac{P_j(\mu)P_l(\mu)}{1-\eta(u)\mu}, \tag{16}$$

where

$$\eta(u)=iy(u)=ik/\sigma_T(u). \tag{17}$$

The set of equations (15) is made finite by assuming that $J_l=0$ for $l>L$.

By virtue of the recursion relations satisfied by the Legendre polynomials, it follows that

$$\frac{1}{\eta}-(2l+1)A_{jl}(\eta)-(l+1)A_{j,l+1}(\eta)-lA_{j,l-1}(\eta)=\frac{\delta_{jl}}{\eta}, \tag{18}$$

with an identical recursion relation holding for the index j because the A_{jl} are symmetric (Eq. (16)). By means of Eq. (18), it is possible to express all the A_{jl} in terms of A_{00}.

Observing that as l becomes infinite, A_{jl} approaches zero, and that

$$A_{00}(\eta)=\eta^{-1}\tanh^{-1}\eta=Q_0(1/\eta)/\eta, \tag{19}$$

it may be shown in general[11] that

$$A_{jl}(\eta)=Q_j(1/\eta)P_l(1/\eta)/\eta, \quad l\leq j, \tag{20}$$

where $Q_j(x)$ is a Legendre function of the second kind.[12]

4. Relation Between the B_L and P_L Procedures. Convergence

From Eqs. (15) and (20), it follows that if $J_l=0$ for $l>L$, then

$$N_{l'}/N_{l'+1}=Q_{l'}(1/\eta)/Q_{l'+1}(1/\eta), \quad l'\geq L. \tag{21}$$

The B_L approximation is equivalent to using the Eqs. (14), but setting

$$N_{L+1}=N_LQ_{L+1}(1/\eta)/Q_L(1/\eta),$$

rather than zero as in the P_L approximation. Such a

[11] I. Waller, Arkiv. Mat. Astron. Fysik **34A**, No. 3 (1947).

[12] See, for example, Jahnke and Emde, *Tables of Functions* (Dover Publications, Inc., New York, 1943), p. 109 ff. Since η is an imaginary quantity, one must be careful to use the correct form of the function, which Jahnke and Emde write as a German script Q.

procedure corresponds to using Wick's Eq. (28),[13] whereas the P_L approximation is obtained if one sets $\lambda_L=0$ in W-20.

An alternative means of comparing the methods is to note that the $L+1$ equations obtained from Eq. (14) by setting $N_{L+1}=0$ may be recast into the form of Eq. (15) with the A_{jl} replaced by A_{jl}', where

$$A_{jl}'=\frac{1}{2}\int_{G_{L+1}} d\mu \frac{P_j(\mu)P_l(\mu)}{1-\eta\mu}, \tag{22}$$

where the subscript G_{L+1} on the integral indicates that the Gaussian quadrature formula[14] with $L+1$ discrete ordinates is to be used to evaluate the integral. Note that $A_{j,L+1}'\equiv 0$, inasmuch as in the Gaussian formula, the integrand is evaluated at the zeros of $P_{L+1}(\mu)$.

This relationship between the B_L and P_L approximations also follows directly from the fact that the Gaussian integration procedure of Wick,[15] in which one considers only the neutron flux in directions whose direction cosines are zeros of $P_{L+1}(\mu)$, is, as may readily be shown, completely equivalent to the P_L approximation.

Since, from Eq. (22), it is clear that the A_{jl}' obey the same recursion relations as do the A_{jl},[15a] the P_L case is obtained from the B_L by simply substituting A_{00}' for A_{00}.[16]

In the P_1 approximation, for example,

$$A_{00}'=(1+y^2/3)^{-1}=1-y^2/3+y^4/9+\cdots, \tag{23a}$$

while

$$A_{00}=y^{-1}\tan^{-1}y=1-y^2/3+y^4/5+\cdots. \tag{23b}$$

Thus, the P_1 and B_1 cases are seen to be equivalent for y^2 small. In general, the earliest disagreement in the expansion of A_{00} and A_{00}' occurs in terms of the order y^{2L+2}.[17]

In estimating the convergence of the P_L or B_L procedures with L, it is convenient to make use of an expansion of the type given in Eq. (11) for the hydrogen integral, J_j^H.

$$J_j^H(k,u)=\sum_{n=0}^{\infty} T_{j0}{}^{Hn}\frac{\partial^n}{\partial u^n}(N_j(k,u)\sigma_s{}^H(u)), \tag{24}$$

where $T_{j0}{}^{Hn}$ is given by Eqs. (10) and (12).

The integral in Eq. (12) may be performed explicitly, yielding

$$T_{j0}{}^{Hn}=\sum_{l=0}^{\lambda}\frac{(-)^{l+n}(2j-2l-1)!}{(j-l-1)!(j-2l)!}\frac{1}{(j-2l+2)^{n+1}}, \tag{25}$$

where $\lambda=j/2$ if j is even and $\frac{1}{2}(j-1)$ if j is odd.

The most important term of Eq. (25) is that for $n=0$. Suppressing, for convenience, the superscript "H," we find,[18]

$$\begin{array}{lll} T_{00}{}^0=1 & T_{30}{}^0=0 & T_{60}{}^0=\frac{1}{64}, \\ T_{10}{}^0=\frac{2}{3} & T_{40}{}^0=-1/24 & \\ T_{20}{}^0=\frac{1}{4} & T_{50}{}^0=0 & \end{array} \tag{26}$$

and, in general

$$T_{2l+1,0}{}^0=0,$$

$$T_{2l,0}{}^0=\frac{(-)^{l+1}(2l-3)!}{2^{2l-2}(l-2)!(l+1)!}\to\left(\frac{e}{l}\right)^{\frac{1}{2}}\frac{(-)^{l+1}}{2(\pi)^{\frac{1}{2}}} \tag{27}$$

as $l\to\infty$.

$l=2, 3, \cdots$.

Thus, the only extra contributions in going from the B_{n-1} to the B_n case, for n odd, come from the derivative terms in Eq. (24). In the P_n case, however, there is an additional improvement in going from P_{n-1} to P_n due to the fact that A_{jl}' is evaluated by the $(n+1)$ point rather than the n-point Gauss formula.

The smallness of the quantities T_{L0} is the justification for believing the B_L approximation to be more accurate than the corresponding P_L approximation. This is because in the B_L case one ignores products of the form $T_{L+1,0}N_{L+1}(k,u)$, whereas in the P_L procedure one neglects N_{L+1} itself.[19] The neglect of N_{L+1} in the P_L approximation manifests itself by the error in the coefficients A_{lj}' in Eq. (22) in addition to the error due to omitting the higher terms in the sum, the latter error being present also in the B_L approximation. It is therefore to be expected that as L increases, the difference between the B_L and P_L approximations becomes less. It is also apparent that the difference between the two approximations becomes larger as y increases due to the inadequate representation of the integral in Eq. (22) by the Gaussian procedure.[20]

---

[13] See G. C. Wick, Phys. Rev. **75**, 738 (1949). We refer to this paper as "W."

[14] See, e.g., J. B. Scarborough, *Numerical Mathematical Analysis* (Johns Hopkins Press, Baltimore, 1950), second edition, p. 145 ff.

[15] G. C. Wick, Z. Physik **121**, 702 (1943).

[15a] This is true only if $j+l<2L+2$; see footnote (17).

[16] For even L, A_{00}' does not approach zero as $k\to\infty$ due to the fact that one of the zeros of $P_{2n+1}(\mu)$ is for $\mu=0$, and $(1-iy\mu)^{-1}=1$ at this point for all k. This behavior of A_{00}' prevents $N_0(k)$ from approaching zero as $k\to\infty$ which introduces the well-known delta function in the spatial distribution of flux calculated by means of a P_{2L} approximation. This difficulty clearly does not arise in the B_L approximation since the correct value of A_{00} is used for all L. For this reason, one never encounters, in practice, a P_{2L} approximation, although a B_{2L} approximation is perfectly legitimate. In fact, the arguments of Sec. II, Part 3 indicate that it is desirable to use even L in the B type approximation.

[17] This is a consequence of the well-known theorem to the effect that a quadrature formula based on the nth Gaussian quadrature approximation integrates rigorously any polynomial of degree up to and including $2n-1$.

[18] The $T_{j0}{}^0$ may be found from Eq. (25), or evaluated directly from the generating function for the Legendre polynomials. Note that the $T_{l0}{}^0$ are identical with the $g_l(0)$ in W.

[19] This argument is particularly pertinent for heavier scatterers, for which the $T_{L0}{}^0$ decrease more rapidly with increasing L.

[20] This point gives the motivation for the use of the B_L approximation in shielding problems, when one is interested in values of η close to unity. See reference 6.

TABLE I. Comparison of B_L and P_L approximations in constant cross section case.

	$y=1$		$y=5$	
L	$\mathfrak{N}_0(B_L)$	$\mathfrak{N}_0(P_L)$	$\mathfrak{N}_0(B_L)$	$\mathfrak{N}_0(P_L)$
0	3.660	∞	0.3787	∞
1	1.660	1	0.2990	0.0400
2	1.831	2.067	0.3218	1.1067
3	"	1.794	"	0.15144
4	1.832	1.839	0.3238	0.5422
5	"	1.832	"	0.2313
6				0.4074
7				0.2781

It should be noted that if one writes[1]

$$N_l(k,u)=\sum_{j=l,l+2,\cdots}\frac{(ik)^j}{j!}N_l^{(j)}(u), \tag{28}$$

both the P_L and B_L approximations will give the correct value of $N_l^{(j)}(u)$ for values of j up to and including $j=2L$. In particular, if $L=1$, both methods give the correct second moment of the neutron flux from a point source, $\bar{r}^2$, which is related to the so-called neutron "age," τ:

$$\bar{r}^2=6\tau=3N_0^{(2)}(u)/N_0^{(0)}(u). \tag{29}$$

It is to be expected that although the B_L approximation does not yield rigorously correct values for $N^{(j)}(u)$ with $j>2L$, the error in these coefficients is less than when the P_L approximation is used. Specific numerical comparison of the two methods for various values of L will be given in the aforementioned subsequent paper.

Another comparison of the convergence of the P_L and B_L approximations can be obtained by considering the integral of the neutron flux with respect to lethargy for the case in which the cross section is independent of energy. Setting

$$\int_{-\infty}^{\infty}N_j(k,u)du=\mathfrak{N}_j(k), \tag{30}$$

and assuming that hydrogen is the only scattering material, and that $\sigma_a^H=0$, one obtains from Eq. (14)[21]

$$\sigma_s^H\mathfrak{N}_j-\frac{ik}{2j+1}[(j+1)\mathfrak{N}_{j+1}+j\mathfrak{N}_{j-1}]=\delta_{j0}+\sigma_s^H T_{j0}^{H0}\mathfrak{N}_j. \tag{31}$$

These equations have been solved for two values of y by the P_L and B_L procedures, the results being shown in Table I. Note the far superior convergence of the B_L procedure. Even for y as large as 5, the B_2 approximation is good to better than one percent, while the B_1 approximation gives a better answer than does the P_7.

The equality of the results for B_{2i} and B_{2i+1} with $i>0$ results from the fact that the $T^{H0}{}_{2j+1,0}=0$. Clearly, all of the derivative terms in Eq. (24) vanish when the lethargy integration [Eq. (30)] is performed. As was mentioned previously, the P_{2i+1} approximation is better than the P_{2i} because one more point is used in the Gauss quadrative formulas for the A_{jl}' in the P_{2i+1} case than in the P_{2i}. The large error in the B_0 approximation for small y arises from the fact that for hydrogen, $\sigma_{Tr}=\frac{1}{3}\sigma$, where σ_{Tr} is the "transport cross section." The B_0 approximation, ignoring as it does J_1, has the effect of replacing σ_{Tr} by σ. Of course the P_0 approximation, which sets $N_1=0$, predicts no net current, and hence no diffusion of neutrons, and is thus meaningless.

[21] Equation (31) is the equivalent of W-15 for $\eta=0$, since Wick's $\gamma_j=1-g_j$ so that $\gamma_l(0)=1-T_{l0}^{H0}$. Wick's η, a variable in Laplace transform space, is not to be confused with our $\eta=ik/\sigma_T$.

III. APPROXIMATE SOLUTIONS FOR THE P_1 AND B_1 CASES

The approximate methods mentioned in Sec. I, i.e., the age diffusion and S.G. methods, involve essentially the use of Eq. (11) in one or both of the simultaneous equations (14) in the P_1 approximation.

In the age-diffusion method, two terms are retained in the expansion of J_0 and only one in the expansion of J_1. This yields the equation

$$(\sigma_a(u)+D(u)k^2)N_0(k,u)=-\sum_\gamma\frac{\partial}{\partial u}[\xi^\gamma\sigma_s^\gamma(u)N_0(u)]+S(u), \tag{32}$$

where

$$D(u)=(3\sigma_{Tr}(u))^{-1}, \tag{33a}$$

$$\sigma_{Tr}(u)=\sigma_T(u)-\sum_\gamma\sum_{L'=0}^{\infty}B_{L'}{}^{\gamma c}(u)T_{1L'}{}^{\gamma 0}=\sigma_T(u)-\sum_\gamma\sigma_s^\gamma\langle\cos\theta^\gamma\rangle_{Av}. \tag{33b}$$

$$\xi^\gamma\sigma_s^\gamma(u)=-\sum_{L'=0}^{\infty}B_{L'}{}^c(u)T_{0L'}{}^1. \tag{33c}$$

For isotropic scattering in the center-of-mass system,

$$\langle\cos\theta^\gamma\rangle_{Av}=\frac{2\gamma}{3}, \tag{34a}$$

and

$$\xi^\gamma=1-\frac{\alpha}{1-\alpha}\ln\frac{1}{\alpha}, \tag{34b}$$

$$\alpha=\left(\frac{\gamma-1}{\gamma+1}\right)^2. \tag{34c}$$

It is clear that the accuracy of this method depends upon the rapid convergence of the series in Eq. (11), which occurs when the lethargy change in a collision, U, is small. This criterion is not satisfied for moderation by hydrogen, since in one collision the neutron can lose all of its energy.

The slowing-down density, $q(k,u)$, is given in this approximation by

$$q(k,u)=\sum_\gamma\xi^\gamma\sigma_s^\gamma(u)N_0(k,u), \tag{35}$$

as may be verified by substituting a two-term Taylor's series expansion in the general expression for the slowing-down density.[3]

Equation (32) may be solved readily yielding

$$N_0(k,u)=(\xi\sigma_s)^{-1}\int_0^u S(u')\times\exp\left(-\int_{u'}^u\frac{\sigma_a(u'')+D(u'')k^2}{\xi\sigma_s(u'')}du''\right)du', \quad (36)$$

where $\xi\sigma_s$ has been written for $\sum_\gamma \xi^\gamma\sigma_s^\gamma$.

In the S.G. approach, the expansion Eq. (11) is made for J_1^γ and for all of the J_0^γ except J_0^H. Again only one term is retained in the J_1^γ series, and two in the expansion for the J_0^γ. One obtains in this manner

$$[\sigma_a(u)+\sigma_s^H(u)+k^2D(u)]N_0(k,u)=J_0^H(u)-\sum_{\gamma\neq 1}\frac{\partial}{\partial u}(\xi^\gamma\sigma_s^\gamma(u)N_0(k,u))+S(u). \quad (37)$$

Equation (37) may be solved easily if one assumes that all of the energy degradation takes place in hydrogen. One thus has $\xi^\gamma=0$, so the slowing-down density is given by

$$q(k,u)=J_0^H(k,u)=\int_0^u e^{-(u-u')}\sigma_s^H(u')N_0(k,u')du', \quad (38)$$

and obeys the equation[22]

$$q(k,u)+\frac{\partial q(k,u)}{\partial u}=\sigma_s^H(u)N_0(k,u). \quad (39)$$

Equations (37) and (39) are simultaneous differential equations for N and q. Solving for q, one finds,

$$q(k,u)=\int_0^u S(u')(1-\rho(u'))\times\exp\left(-\int_{u'}^u\rho(u'')du''\right)du', \quad (40)$$

where

$$\rho(u)=\frac{\sigma_a(u)+k^2D(u)}{\sigma_a(u)+k^2D(u)+\sigma_s^H(u)}. \quad (41)$$

Note that the S.G. result tends to be smaller than the age-diffusion result due to the factor $(1-\rho)$ multiplying S. However, there is a compensating factor which tends to makes the S.G. result larger, namely the fact that $\rho(u)$ is smaller than the corresponding quantity in the exponent of Eq. (36).

If the same approximations as the two discussed above had been made on the B_1 rather than the P_1 approximation, one would obtain "modified" age-diffusion and S.G. methods. These modified methods may be put in the same form as the more usual methods, if the transport cross section is redefined:

$$\frac{1}{3\sigma_{Tr}}\rightarrow\frac{1}{3\text{“}\sigma_{Tr}\text{”}}=\text{“}D\text{”}=\frac{1-A_{00}(y)}{y^2[A_{00}(y)\sigma_T(u)-3A_{11}(y)\langle\cos\theta\rangle_{Av}\sigma_s(u)]}. \quad (42)$$

Here A_{00} and A_{11} are defined by Eq. (20). Note that as $y\rightarrow 0$, "σ_{Tr}"$\rightarrow\sigma_{Tr}$.

The same redefinition of σ_{Tr} will serve to convert the so-called Greuling-Goertzel method[23] into a "modified" method. This modification is of value in treating a situation in which the Taylor's series expansion of the lethargy integrals are justified, but the use of the P_1 approximation is not.

Expanding "σ_{Tr}" in powers of y, we find,

$$\text{“}\sigma_{Tr}\text{”}=\sigma_{Tr}+\sigma_s\left(\frac{4}{15}y^2+\frac{9}{25}y^4+\cdots\right), \quad (43)$$

so that if σ_{Tr} and $y^2\sigma_s$ are of the same order of magnitude, the correction is important.

IV. THE NEUTRON AGE

The so-called neutron age, defined in Eq. (29), is of interest for several reasons. First, it is a quantity which enters directly into the calculation of reactor criticality and neutron flux distributions.[24] Secondly, it can be and has been measured,[25] providing to some extent a comparison between theory and experiment. Finally, since the correct value of the age is given in the P_1 approximation (see Sec. II), it may be obtained by a relatively simple calculation. In particular, there is no need to invert the Fourier transform of the flux as would be necessary if the experimental and theoretical flux distributions from a point source were to be compared. A study of the expressions for the age also gives some insight of the relationships among the various approximate treatments.

There is a slight difference between the theoretical expressions for and the numerical values of the age depending upon whether one is interested in the ratio of the second to the zeroth moments of the flux or the corresponding quantity for the slowing-down density. We refer to these as τ_N and τ_q, respectively. The quantity τ_N is given by Eq. (29), while for mixtures of hydrogen with heavy moderators which do not slow

[22] Of course, J_0^H obeys Eq. (39) whether or not the assumption is made that only hydrogen slows down. In such a case, one still solves Eqs. (37) and (39) simultaneously, but the terms involving ξ^γ in Eq. (37) do not drop out.

[23] E. Greuling, Phys. Rev. 87, 177 (1952). Clark, Goertzel, and Greuling (unpublished).

[24] See references 2 and 3 for applications.

[25] Hill, Roberts, and Fitch, U. S. Atomic Energy Commission report AECD-3392 (unpublished).

down τ_q is given by[26]

$$\tau_q=\frac{1}{2}\frac{J_0{}^{H(2)}(u)}{J_0{}^{H(0)}(u)}, \tag{44a}$$

where

$$J_0{}^{H(i)}(u)=\int_0^u e^{-(u-u')}\sigma_s{}^H(u')N_0{}^{(i)}(u')du'. \tag{44b}$$

The quantity τ_N is larger than τ_q at the same lethargy by an amount sometimes referred to as the "last-flight correction." Experimental measurements of the age by means of the activation of a thin resonance absorber (e.g., indium) give τ_N, whereas τ_q is somewhat more appropriate to reactor calculations since the "last-flight" correction is in a sense included in the treatment of the thermal group.[27]

Continuing to consider mixtures of hydrogen with nonslowing-down heavier moderators, we can write expressions for the quantities appearing in Eq. (28) and (44a)[28]:

$$\tfrac{1}{2}N_0{}^{(2)}=[\sigma_a(u)+\sigma_s{}^H(u)]^{-1}\times\left[N_1{}^{(1)}(u)+\int_0^u(1-s(u'))N_1{}^{(1)}(u')\times\exp\left(-\int_{u'}^u s(u'')du''\right)du'\right], \tag{45a}$$

with

$$N_1{}^{(1)}(u)=\tfrac{1}{3}[\sigma_{Tr}(u)+\tfrac{2}{3}\sigma_s{}^H(u)]^{-1}\times\left[N_0{}^{(0)}(u)+\frac{3}{2}\int_0^u(1-\lambda(u'))N_0{}^{(0)}(u')\times\exp\left(-\frac{3}{2}\int_{u'}^u\lambda(u'')du''\right)du'\right], \tag{45b}$$

and

$$N_0{}^{(0)}(u)=[\sigma_a(u)+\sigma_s{}^H(u)]^{-1}\times\left[S(u)+\int_0^u(1-s(u'))S(u')\times\exp\left(-\int_{u'}^u s(u'')du''\right)du'\right]. \tag{45c}$$

Here

$$s(u)=\frac{\sigma_a(u)}{\sigma_a(u)+\sigma_s{}^H(u)}, \tag{46a}$$

and

$$\lambda(u)=\frac{\sigma_{Tr}(u)}{\sigma_{Tr}(u)+\frac{2}{3}\sigma_s{}^H(u)}. \tag{46b}$$

These equations determine τ_N. Equations (45) are derived by inserting the expansion Eq. (28) in Eq. (14), and equating coefficients of like powers of k. Three simultaneous integral equations are obtained for $N_0{}^{(0)}$, $N_0{}^{(2)}$, and $N_1{}^{(1)}$, which are readily solved when the approximation $\xi^\gamma=0$ is made.

To obtain τ_q, note that with the above physical assumptions Eq. (14) yields:

$$\tfrac{1}{2}J_0{}^{H(2)}=\tfrac{1}{2}(\sigma_a+\sigma_s{}^H)N_0{}^{(2)}-N_1{}^{(1)}=\int_0^u(1-s(u'))\times\exp\left(-\int_{u'}^u s(u'')du''\right)N_1{}^{(1)}(u')du', \tag{47}$$

and

$$J_0{}^{H(0)}=(\sigma_a+\sigma_s{}^H)N_0{}^{(0)}-S=\int_0^u S(u')(1-s(u'))\times\exp\left(-\int_{u'}^u s(u'')du''\right)du'. \tag{48}$$

Thus, τ_q is given by Eq. (44a) in conjunction with Eqs. (47) and (48).

The quantities occurring in the expressions for τ_N and τ_q are simpler when there is no absorption, for then $s(u)=0$, and we have

$$\tfrac{1}{2}N_0{}^{(2)}(u)=(\sigma_s{}^H(u))^{-1}\times\left(N_1{}^{(1)}(u)+\int_0^u N_1{}^{(1)}(u')du'\right), \tag{49a}$$

$$\tfrac{1}{2}J_0{}^{H(2)}(u)=\int_0^u N_1{}^{(1)}(u')du', \tag{49b}$$

with $N_1{}^{(1)}(u)$ given still by Eq. (45b) and

$$N_0{}^{(0)}(u)=T(u)/\sigma_s{}^H(u), \tag{49c}$$

$$T(u)=\int_0^u S(u')du'+S(u). \tag{49d}$$

Of course, $J_0{}^{H(0)}(u)$ is given by Eq. (47b) as before.

The Selengut-Goertzel approximation gives $J_0{}^{H(0)}$ and $N_0{}^{(0)}$ correctly, and also predicts the correct de-

[26] In general, one would add to numerator and denominator contributions to q from the heavier moderators, which would be written $\frac{1}{2}\sum_\gamma \xi^\gamma\sigma_s{}^\gamma N_0{}^{(2)}$ and $\sum_\gamma \xi^\gamma\sigma_s{}^\gamma N_0{}^{(0)}$, respectively.

[27] This does not yield a rigorous description of the transition from the slowing-down region to the thermal energy region, since the attainment of thermal equilibrium is actually a complex phenomenon depending upon, among other things, the molecular structure of the moderating materials.

[28] Similar expressions, in various approximations, have been given by Marshak (reference 1 and H. C. Volkin, J. Appl. Phys. 25, 83 (1954).

pendence of $N_0^{(2)}$ or $J_0{}^{H(2)}$ on $N_1^{(1)}$. However, we have a new expression for the last quantity:

$$N_1^{(1)}(u)=\tfrac{1}{3}N_0^{(0)}(u)/\sigma_{Tr}(u). \tag{50}$$

This change is equivalent to making the following replacement in Eq. (45b):

$$\exp\left(-\frac{3}{2}\int_{u'}^{u}\lambda(u'')du''\right)\rightarrow\tfrac{2}{3}\delta(u-u')/\lambda(u'). \tag{51}$$

In the age-diffusion approximation as applied to hydrogen moderation with $\sigma_a=0$,[29] there is, in addition to the approximation of Eq. (50) the additional approximation of replacing Eq. (49c) by

$$N_0^{(0)}(u)=(\sigma_s{}^H(u))^{-1}\int_0^u S(u')du'. \tag{49c'}$$

Hence in the age-diffusion approximation for a monoenergetic source at $u=0$,[30]

$$\tau_{AD}=\frac{1}{3}\int_0^u\frac{1}{\sigma_{Tr}(u')\sigma_s{}^H(u')}du'. \tag{52}$$

The S.G. approximation adds a "S.G. first-flight correction" to the age-diffusion result, to wit

$$\tau_{qSG}=\tau_{AD}+\tfrac{1}{3}(\sigma_{Tr}(u)\sigma_s{}^H(u))^{-1}{}_{u=0}. \tag{53}$$

For $\sigma_a=0$ and a monoenergetic source at $u=0$, the correct expression for $N_1^{(1)}$ (still assuming $\xi^\gamma=0$) is given by:

$$\begin{aligned}N_1^{(1)}(u)=&\tfrac{1}{3}(\lambda(u)/\sigma_{Tr}(u))\\&\times\left[\frac{1+\delta(u)}{\sigma_s{}^H(u)}+\frac{3}{2}\frac{(1-\lambda(0))}{\sigma_s{}^H(0)}\right.\\&\times\exp\left(-\frac{3}{2}\int_0^u\lambda(u')du'\right)+\frac{3}{2}\int_0^u\frac{(1-\lambda(u'))}{\sigma_s{}^H(u')}\\&\left.\times\exp\left(-\frac{3}{2}\int_{u'}^u\lambda(u'')du''\right)du'\right].\end{aligned} \tag{54}$$

If this expression is inserted into Eq. (49b) [note that $J_{0q}{}^{(0)}=1$ for $\sigma_a=0$ and $S(u)=\delta(u)$] there is a contribution to the age from the delta-function term above which gives the correct transport theory first-flight correction, F.F.C.,

$$\text{F.F.C.}=\lambda(u)/3\sigma_s{}^H(u)\sigma_{Tr}(u)]_{u=0}. \tag{55}$$

This is smaller than the S.G. expression for F.F.C. because the diffusion theory σ_{Tr} is replaced by

$$\sigma_{Tr}{}'=\sigma_{Tr}/\lambda=\sigma_{Tr}+\tfrac{2}{3}\sigma_s{}^H. \tag{56}$$

The added term takes account of the fact that a hydrogen scattering event removes neutrons for the initial energy range, and thus has the effect of an absorption.[31] Hence the total hydrogen scattering cross section, rather than $\sigma_s{}^H(1-\langle\cos\theta\rangle_{Av})=\tfrac{1}{3}\sigma_s{}^H$ appears in σ_{Tr}.

From Eq. (51), it might be expected that the other terms in the correct age expression would tend to make up the difference between the rigorous expression for F.F.C. and the S.G. expression for F.F.C. Although this is to some extent true for λ (i.e., $\sigma_{Tr}/\sigma_{Tr}{}'$) greater than 2/3, there is only partial compensation. For example, with energy-independent cross sections and a monoenergetic source,

$$\begin{aligned}\tau_q=&\tau_{AD}+\tfrac{1}{3}(\sigma_{Tr}{}'\sigma_s{}^H)^{-1}\\&\times\left[1-\lambda^{-1}(1-\lambda)\left(1-\frac{2}{3\lambda}\right)(1-e^{-\frac{3}{2}\lambda u})\right].\end{aligned} \tag{57}$$

If the factor $[1-(2/3\lambda)]$ in the second term under the brackets were absent, then for large u the "effective" first-flight collision would equal the S.G. first-flight correction. Because of this factor [which is associated with the inadequacy of Eq. (51)] the extra terms which occur in the transport theory expression in addition to F.F.C. make up only one third of the difference between the S.G. and transport theory F.F.C. even if λ is close to unity. (The case $\lambda\sim1$ corresponds to the presence of substantial amounts of nonhydrogenous scattering material, and in this case, the error in the S.G. age is small.)

For $\lambda<\frac{2}{3}$, as is the case for pure hydrogen (then $\lambda=\frac{1}{3}$) the overestimate of the age by the S.G. method is greater than one would expect by comparing the S.G. and transport theory first-flight corrections. In fact, it may be seen from Eq. (57) that for constant cross sections if $\lambda<\frac{2}{3}$, the correct τ_q is actually less than τ_{AD}.

Some of these conclusions also apply to the case of variable cross sections, as may be seen from the numerical results discussed in the next section.

V. NUMERICAL CALCULATIONS

The numerical solution of the B_L and P_L approximations for $L=1$, 2, and 3, as well as of the S.G. and "modified" S.G. methods, is at present under way on the IBM facilities at the Knolls Atomic Power Laboratory as well as the New York University UNIVAC. The procedures followed and the numerical results obtained will be the topic of a subsequent paper. The inversion

[29] Actually, the "age" can be written in the age-diffusion approximation for an arbitrary mixture of moderators. The quantities entering into the age are given by

$$\tfrac{1}{3}N_0^{(2)}(u)=\tfrac{1}{3}(\xi\sigma_s(u))^{-1}\int_0^u\frac{N_0^{(0)}(u')}{\sigma_{Tr}(u')}\exp\left(-\int_{u'}^u\frac{\sigma_a(u'')}{\xi\sigma_s(u'')}du''\right)du',$$

$$N_0^{(0)}=\tfrac{1}{3}(\xi\sigma_s(u))^{-1}\int_0^\infty S(u')\exp\left(-\int_{u'}^u\frac{\sigma_a(u'')}{\xi\sigma_s(u'')}du''\right)du',$$

where we have written $\xi\sigma_s$ for $\Sigma_\gamma(\xi\sigma_s)^\gamma$. Note that no "last collision correction" is predicted by this theory; thus, $\tau_{qAD}=\tau_{NAD}$.

[30] This expression is also valid for a polyenergetic source at energies below the minimum source energy.

[31] This has been pointed out by L. C. Biedenharn and T. Welton, private communication to M. Shapiro.

TABLE II. τ_N for 1.4 ev neutrons in water.

Ref.	Accurate	S.G.	A.D.
a	25.3 cm²	30.9 cm²	
b	23.6	28.8	
c	26.8	32.8	17.5 cm²
d	25.3		

a No approximations.
b Oxygen scattering assumed isotropic.
c Oxygen assumed not to slow down.
d Approximations b and c both made.

of the Fourier transform of the flux will be discussed also, so that comparison may be made with measured spatial flux distributions.

As a by-product of the P_1 (or B_1) calculations, the neutron age is obtained, since, as is seen from Eq. (29),

$$\tau_N=\tfrac{1}{2}\lim_{k\to 0}\frac{\partial^2}{\partial k^2}N_0(k,u)/\lim_{k\to 0}N_0(k,u), \qquad (58a)$$

while, analogously,

$$\tau_q=\tfrac{1}{2}\lim_{k\to 0}\frac{\partial^2}{\partial k^2}q(k,u)/\lim_{k\to 0}q(k,u). \qquad (58b)$$

These equations are of course identical with Eqs. (29) and (44d), and the analytical expressions given in the previous section for the quantities entering into the expressions for τ are identical with Eqs. (58) so long as the same physical assumptions are made in both cases.

The ages of indium resonance neutrons (1.44 ev) slowing down in water as calculated by the various methods are given in Table II. The calculations have been done considering and ignoring energy degradation by oxygen in order to indicate the effect of the physical assumptions made in Sec. IV. Also, the calculations have been done both with and without the assumption that the oxygen scattering is isotropic in the center of mass system [see Eqs. (33) and (34)]. The oxygen scattering cross sections and the expansion coefficients were obtained from Goldstein.[32] The neutron source was assumed to be a fission spectrum, given by the Watt formula.[33]

The experimental value for τ at 1.44 ev is 30.5 cm².[25] Note that the S.G. method gives an answer which compares well with experiment, while the rigorous expression is about 16% low. Since it was pointed out in the previous section that the S.G. method overestimates τ, the agreement between the calculated S.G. and experimental ages must be regarded as fortuitous. The ages calculated here agree with the calculations of the second spatial moment of N_0 by Certaine and Aronson.[34]

The disagreement between calculated and experimental ages is difficult to understand, but numerical calculations have shown that slight errors in the source-energy spectrum can lead to large errors in τ, since it is possible that too many neutrons may have been put in at energies at which the cross section is low, i.e., "windows," which leads to a first-flight correction which is too large. It is felt that more experimental determinations of τ should be performed, possibly with a source of neutrons whose energy distribution is better known than is the fission spectrum.

From Table II, one notes that the errors introduced by ignoring the anisotropy of the oxygen scattering and the slowing-down properties of oxygen are in opposite directions, and tend to cancel each other out. Thus, in a practical calculation, one may justifiably ignore both effects and expect to get a reasonably good answer.

VI. ACKNOWLEDGMENTS

The authors are grateful for opportunities to discuss and compare results with members of several other laboratories with interests in this field, particularly Nuclear Development Associates; Westinghouse Atomic Power Division; NACA, Cleveland; and GE Aircraft Nuclear Propulsion Department. We also would like to acknowledge illuminating discussions with Professor E. Greuling.

[32] H. Goldstein, Nuclear Development Associates, Inc., Memo 15C-15 (unpublished).

[33] B. E. Watt, Phys. Rev. **87**, 1037 (1952).

[34] J. Certaine and R. Aronson, Atomic Energy Commission report NDA-15C-40 (unpublished). Similar results have been obtained at several other laboratories.

Tracer Diffusion of Iron in Stainless Steel

V. LINNENBOM, M. TETENBAUM, AND C. CHEEK
Naval Research Laboratory, Washington, D. C.

(Received December 17, 1954)

The tracer diffusion of iron in 18–8 stainless steel has been measured over a wide temperature range by the surface activity decrease method, using radioactive Fe^{55}. The results show that the lattice diffusion coefficients can be described by the equation $D=0.58\exp(-67\,100/RT)$ cm² sec⁻¹. The effect of variations in grain size has been determined, and it has been shown that when the grain size is small, grain boundary diffusion contributes appreciably to the over-all diffusion process, this effect decreasing with increasing temperature.

INTRODUCTION

THE determination of diffusion rates in metal alloys by chemical methods of analysis is based on the existence of measurable concentration differences in the system being investigated and is therefore subject to variations caused by these concentration gradients. The concentration dependency of the diffusion coefficient and the phenomena associated with the Kirkendall effect (i.e., mass flow and resultant lattice distortion), are among the more important of the factors which can affect chemically determined diffusion rates. It is probable that such variables are the cause of the many discrepancies which exist between the results of different investigators. Diffusion data at different temperatures are best described by the exponential equation

$$D=D_0\exp(Q/RT), \tag{1}$$

where D is the measured diffusion coefficient at temperature T, D_0 is a frequency factor whose theoretical significance is not yet completely understood, and Q is the activation energy. If one examines the reported values of D_0 and Q for chemically determined diffusion rates,[1] the most striking feature is the wide range covered, D_0 varying from 10^{-9} to 10 cm²/sec, while Q covers a range from 8 to 140 kcal/mole. Such values, if correct, present a formidable obstacle to any theoretical interpretation of D_0 and Q.

An examination of the data for self-diffusion in pure metals reveals much closer agreement. The reported values of D_0, for example, extend from 10^{-1} to 10 cm²/sec, a much smaller range than that just mentioned. This is undoubtedly because the disturbing effects caused by concentration differences are absent in the case of self-diffusion; the driving force for the latter process involves only the entropy of mixing, and such variables as concentration gradients do not exist. For this reason, theoretical attempts to understand the significance of D_0 have met with the first degree of success when applied to self-diffusion.[2]

In the case of diffusion in alloys a similar advantage results from the use of tracer methods in that the alloy remains chemically homogeneous throughout the course of the diffusion process; the concentration effects which complicate the interpretation of results based on chemical methods are entirely absent. For example, Sonder, Slifkin, and Tomizuka[3] have measured the diffusion of tracer antimony into silver; their results disagree with those of Seith and Peretti,[4] who measured the diffusion of Sb from a 2 percent alloy into pure silver, where an appreciable concentration gradient existed. From the standpoint of theory, tracer diffusion in alloys has also produced promising results; thus Johnson's[5] values for the tracer diffusion coefficients of silver and gold in a 50:50 Ag—Au alloy agree with the theory of Darken[6] relating the over-all interdiffusion coefficient in the alloy and the individual self-diffusion coefficients in the pure metals. LeClaire[2] has pointed out that the work of Ruder and Birchenall[7] on the tracer diffusion of cobalt into nickel, and that of Martin and Asaro[8] on the tracer diffusion of gold into copper, give values for D_0 which agree reasonably well with theory, in contrast to other values for D_0 in the same systems measured with appreciable concentration gradients.

Because of the usefulness of data on tracer diffusion in alloys, it was considered worthwhile to present these results on the tracer diffusion of iron in 18–8 (type 347) stainless steel.

METHOD

The measurement of the diffusion rates in this investigation was accomplished by the surface activity decrease method which is based on the decrease in surface activity as the active tracer diffuses into the metal owing to absorption of radiation by the thickness of metal penetrated by the tracer. If a thin layer of active atoms is deposited on the surface of a comparatively thick sample, the diffusion may be considered to take place from an infinitely thin film into a semi-infinite medium. With these boundary conditions, a

[1] For a summary of the data on diffusion in metals, see C. J. Smithells, *Metals Reference Book* (Interscience Publishers, Inc., New York, 1949), pp. 390–412.
[2] A. D. LeClaire, *Progress in Metal Physics*, edited by B. Chalmers (Interscience Publishers, Inc., New York, 1953), pp. 265–332.
[3] Sonder, Slifkin, and Tomizuka, Phys. Rev. **93**, 970 (1954).
[4] W. Seith and E. Peretti, Z. Elektrochem. **42**, 570 (1926).
[5] W. A. Johnson, Trans. Am. Inst. Met. Engrs. **147**, 331 (1942).
[6] L. S. Darken, Trans. Am. Inst. Met. Engrs. **175**, 184 (1948).
[7] R. C. Ruder and C. E. Birchenall, J. Metals **191**, 142 (1951).
[8] A. B. Martin and F. Asaro, Phys. Rev. **80**, 123 (1950).

solution[9] of Fick's second diffusion law gives

$$C_x = C_0(\pi Dt)^{-\frac{1}{2}} \exp(-x^2/4Dt), \quad (2)$$

where C_x represents the concentration of tracer at distance x in the medium after diffusion time t, and C_0 the total amount of tracer present, defined by $C_0 = \int_0^\infty C_x dx$. Replacing concentrations by activities, and combining Eq. (2) with the exponential law for absorption of radiation as written in Eq. (3),

$$A_{\text{surface}} = \int_0^\infty A_x e^{-\mu x} dx, \quad (3)$$

the following expression can be derived:

$$A_s/A_0 = F = [1 - \text{erf}\{(\mu^2 Dt)^{\frac{1}{2}}\}] \exp(\mu^2 Dt), \quad (4)$$

where A_0 = surface activity measured before diffusion, A_s = surface activity measured after diffusion, μ = absorption coefficient of the radiation in the diffusion medium, D = diffusion coefficient, and t = time. This equation was first used in diffusion work by Steigman, Shockley, and Nix[10] to measure self-diffusion in pure copper.

The calculation of D is carried out most conveniently by a graphical method. A curve is constructed by plotting calculated values of F for various assumed values of the product $\mu^2 Dt$. For an experimentally determined value of F from the measured ratio A_s/A_0, the corresponding value for $\mu^2 Dt$ is then read off the curve. Knowing μ and t, D is then calculated.

Equation (3) is based on an exponential decrease in measured activity with increasing thickness of absorber. For β emitters this exponential absorption law does not necessarily hold true, due to scattering effects and the continuous energy spectrum of the β particles. It is therefore necessary to use an empirical absorption curve which must be measured under the same geometrical conditions as those which exist during the actual diffusion run. X-rays, however, follow the exponential absorption law as required by Eq. (3). For this reason, Fe^{55} was used in this work, rather than the more commonly employed Fe^{59}.

Iron-55, which has a half-life of 2.9 years, decays by K-capture to Mn^{55}, and is identified by the 6 kev x-ray emitted by the latter. The absorption coefficient of this x-ray in 347 stainless steel was calculated from the absorption coefficients for the individual components of the steel according to the formula:

$$\mu_{347} = \rho_{347} \sum_i f_i(\mu/\rho)_i, \quad (5)$$

where μ = linear absorption coefficient (cm^{-1}), ρ = density (g/cc), and f_i = wt fraction of component i in the steel. For 347 stainless steel, the elements considered in calculating μ_{347} were Fe, Cr, Ni, Nb, and Mn.

[9] R. M. Barrer, *Diffusion In and Through Solids* (The Macmillan Company, New York, 1952), p. 45.
[10] Steigman, Shockley, and Nix, Phys. Rev. **56**, 13 (1939).

EXPERIMENTAL

The radioactive Fe^{55} was obtained from neutron bombarded pure iron foil, which had aged sufficiently long (>18 months) to allow essentially complete decay of Fe^{59}. No other activities were found to be present after this time. Activity measurements were carried out by means of a thin end-window, argon-alcohol filled G-M tube, using a conventional scaler and protective lead counting chamber. In measuring A_0 and A_s, a lead ring was placed over the sample so that its outer periphery was shielded during counting, thus eliminating edge effects. Counting times were made sufficiently long so that the statistical counting error was less than 1 percent. The day-to-day performance of the counting equipment was always checked by means of a radioactive standard.

The stainless steel samples used in the diffusion runs were 1-in. diameter, $\frac{3}{8}$-in. thick disks machined from round bar stock. Analysis of the steel showed 68.06 percent Fe, 17.54 percent Cr, 11.33 percent Ni, 1.30 percent Mn, 0.85 percent Nb, 0.70 percent Si, 0.10 percent Mo, and 0.045 percent C. The disks were first given a prediffusion anneal at 1300°C to promote grain growth and thus reduce the effect of grain boundary diffusion (see Discussion). One face of the disk was then hand-polished to a 3–0 finish, followed by successive washes with soap solution, distilled water, and acetone. A quick etch with 20 percent HCl was then used to remove any last traces of oxide. From this point on, all exposure to the air was avoided until the active iron had been deposited. This eliminated any oxide film between the active iron and the steel surface which could conceivably act as a barrier to the diffusion process.

The iron tracer was electroplated onto the prepared steel surface from a saturated ammonium oxalate solution.[11] After plating, the samples were rinsed with distilled water, then acetone, allowed to air-dry, and counted. The plating procedure resulted in a firmly adherent, shiny metallic film, about 5×10^{-5} cm in thickness. Failure to remove completely the oxide film from the steel before plating resulted in a poorly adherent layer of active iron, which could be rubbed off very easily. However, if precautions were taken to avoid contact with air during the short interval between the acid etch and the electro-deposition process, this could be avoided and a firmly adherent layer of active iron was obtained.

The absorption coefficients for iron, chromium, and nickel were measured experimentally, those for niobium and manganese were calculated. For niobium, the absorption coefficient was calculated from a formula given by Victoreen[12]; it was found to contribute about 4 percent to the total absorption by the 347 stainless steel. That for manganese, about 1 percent of the total

[11] R. E. Peterson, Anal. Chem. **24**, 1850 (1952).
[12] J. A. Victoreen, J. Appl. Phys. **20**, 1141 (1949).

absorption effect, was estimated by interpolation between the experimental values obtained for Cr and Fe. The absorption effect due to iron was measured by plating a thin layer of active iron (10^{-4} cm thick) onto a polished copper disk* and measuring its activity I_0. A layer of inactive iron of about the same thickness was then plated over the active iron layer, and the activity I was then determined. The thickness t of the inactive layer was calculated from the weight (determined with an accuracy of ± 0.2 percent) of a known area of plated iron. The absorption coefficient μ was then calculated from the usual exponential formula for absorption,

$$I = I_0 e^{-\mu t}. \tag{6}$$

The absorption coefficient for pure nickel was determined in a similar manner, by plating inactive nickel over the active film. In the case of chromium, a layer of nickel was first plated over the active iron, followed by a layer of chromium; this was necessary to insure a good plate of chromium.

The furnaces for both high-temperature and low-temperature work were similar in design but of different materials. Both consisted of a gas-tight inner lining enclosed in a conventional resistance wire wound alundum tube, adequately insulated with a combination of alumina and magnesia. For high-temperature work, a platinum wound alundum tube was used for heating an inner tube of Sillimanite; for temperatures lower than 1100°C, a Kanthal wire-wound tube was used for heating an inner tube of quartz.

An atmosphere of helium was maintained at a positive pressure throughout the runs by flowing helium into the furnace and allowing it to exit through a bubbler at a rate sufficiently slow to avoid temperature fluctuations within the furnace. With the gas flowing, a constant temperature zone slightly over two inches long existed near the center of the furnace. The helium was purified in an all-glass purification train by flowing it in turn through ascarite, finely divided copper maintained at 550°C, Drierite, magnesium perchlorate, and finally a liquid nitrogen trap. This was successful in preventing any oxidation of the steel samples at elevated temperatures.

In making a diffusion run, the sample was introduced into the hot zone of the furnace after it had reached constant temperature, thus reducing to a minimum the length of time required for the sample to reach temperature. Upon the completion of a run, the sample was withdrawn into the cool zone of the furnace and allowed to cool in an atmosphere of helium before it was exposed to the air.

The temperature in all the runs was held to within $\pm 2°$ by means of voltage regulators and occasional slight adjustments of the applied voltage by manually operated variacs. Both chromel-alumel and platinum-13 percent rhodium-platinum thermocouples were used to measure temperatures.

At the temperatures employed, it has been reported that iron has a high enough vapor pressure to cause possible loss of active iron by evaporation.[13] If this occurred, erroneously high diffusion coefficients would result. To minimize this effect, the samples during the diffusion runs were placed, active face down, on an optically flat quartz disk. To measure the magnitude of this effect, two plated steel samples were run simultaneously at 1000°C. One sample was placed face down on the quartz disk as usual; the other was placed on top of the first sample with active face up and exposed so that evaporation of Fe could take place freely. Activity measurements before and after on both samples showed the same fractional decrease F in surface activity, indicating negligible evaporation effect. At no time during any of the diffusion runs did contamination of the quartz disk occur, although it was in direct contact with the active face of the sample. Nor was any activity ever found in any parts of the furnace. It was therefore concluded that any loss of active Fe by evaporation was small enough so that it could be disregarded.

TABLE I. Absorption coefficients for Mn K_α x-radiation.

Absorbing medium	Fe	Cr	Ni	Nb	Mn
μ(calculated) cm^{-1}	712[a]	450[b]	997[a]	4027[b]	605[c]
μ(measured) cm^{-1}	723	556	1006	...	...

[a] Obtained by interpolation from existing data.
[b] Calculated from Victoreen formula.
[c] Obtained by interpolation between measured values for Fe and Cr.

Three samples of stainless steel were used in obtaining the diffusion data. Two of these samples were preannealed at 1300°C in order to obtain large grain size, while the third was given alternate treatments of cold working and high-temperature annealing in an effort to obtain exaggerated grain size. However, no significant difference in grain size between the three samples was obtained; microscopic examination of all three showed an average grain size of approximately 2–4 mm.

RESULTS

The results of the individual absorption measurements are shown in Table I. The composite absorption coefficient for the steel, calculated from Eq. (5), was found to be 746 cm^{-1}, which is not greatly different from that for pure iron.

The results of the diffusion runs are shown in Table II.

The data in Table II are shown in Fig. 1, in which log D is plotted against $1/T$. An error of ± 10 percent in D was estimated in making this plot; this was based on an evaluation of the errors involved in counting, in temperature measurements, in the measurement of

* Use of a copper disk was permissible because of negligible backscattering in the case of electromagnetic radiation. For β emitters backscattering effects would be appreciable, and steel would have to be used to duplicate more closely the geometry of the actual diffusion experiments.

[13] Marshall, Dornte, and Norton, J. Am. Chem. Soc. **59**, 1161 (1937).

TABLE II. Tracer diffusion of iron in 347 stainless steel.

Temp. °C	Time hr	F	D (cm²/sec) $\times 10^{12}$	$1/T$ (°K⁻¹) $\times 10^4$
808	888	0.775	0.034	9.25
914	311	0.638	0.33	8.42
926	163	0.718	0.36	8.34
948	234	0.597	0.60	8.19
994	100	0.557	1.9	7.89
1005	75	0.592	2.0	7.83
1034	75	0.472	4.5	7.65
1078	47	0.409	12.1	7.40
1129	24	0.421	22	7.13
1181	6	0.483	55	6.88
1200	6	0.465	63	6.79

absorption coefficients, and on an estimate of the error owing to the finite thickness of the electroplated layer of active iron. If the result obtained at the lowest temperature used in these experiments (808°C) is excluded, a straight line is obtained. Values of D_0 and Q in Eq. (1) were obtained by a least-squares analysis of the data in Table II, omitting the results at 808°C. This gives for the tracer diffusion of iron in stainless steel

$$D=0.58\exp(-67\,100/RT). \quad (7)$$

The effect of grain size on the diffusion rate is shown in Table III. An initially fine-grained steel sample (<0.01 mm) was used for a series of successive diffusion runs at 995°C. As grain growth occurred during the successive diffusion anneals at this temperature, the measured value of D decreased until a limiting value of 1.9×10^{-12} cm²/sec was reached. It might be pointed out that at 995°C this limiting value was obtained at a much smaller grain size than that which existed in the samples used for the diffusion runs. The larger grain size of the latter, however, proved necessary at the lower temperatures in order to eliminate grain boundary effects.

DISCUSSION

It is believed that Eq. (7) is representative of lattice diffusion. The high value for D at 808°C is probably due to the increasing influence of grain boundary diffusion at lower temperatures. Since the activation energy for this latter process is considerably less than for lattice diffusion, the relative effect of the grain boundary diffusion process should increase with decreasing temperature. The work of Hoffman and Turnbull[14] on self-diffusion in silver is illustrative of this effect. The same phenomenon was noticed by Sonder, Slifkin, and Tomizuka during their investigation of the diffusion of tracer antimony into silver.[3]

TABLE III. Variation of D with grain size at 995°C.

Time (hr)	Grain size (mm)	$D(\times10^{12})$ cm²/sec
140	<0.01	5.5
120	0.02−0.1	3.7
96	0.1 −0.2	2.3
139	0.2 −0.3	1.9
114	0.2 −0.3	1.9

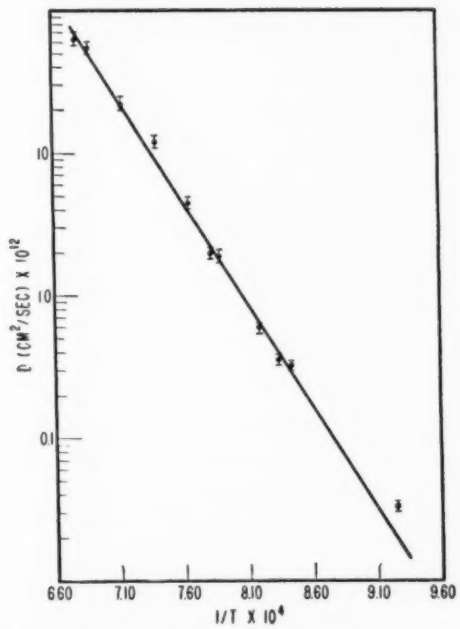


FIG. 1. Tracer diffusion of iron in stainless steel.

The relative contributions of the two diffusion mechanisms to the measured diffusion will also depend on the number of paths available for each process. As the grain size is increased at a given temperature, the number of paths available for boundary diffusion become less and less, until some point is reached where further grain growth produces no change in the measured diffusion coefficient. At this point, grain boundary effects have become negligible.

For each polycrystalline metal, therefore, it would appear that there is a temperature below which grain boundary diffusion gradually becomes predominant with decreasing temperature, but above which it becomes negligible, and the measured diffusion coefficients become identical with the lattice diffusion coefficients. This temperature is a function of grain size and grain orientation, which fix the paths available for boundary diffusion. For the steel samples used in these experiments (2–4 mm average grain size), this temperature appears to be somewhat greater than 800°C; for finer-grained samples, it would be higher. This is supported by the results shown in Table III. The initially fine-grained steel gave a measured D at 995°C of 5.5×10^{-12} cm²/sec, as compared to a D of 1.9×10^{-12} cm²/sec for the large-grain samples. This is evidently due to a boundary diffusion effect, since as grain growth occurred during the successive diffusion anneals at 995°C on this particular sample, the diffusion rate decreased

[14] R. E. Hoffman and D. Turnbull, J. Appl. Phys. **22**, 634 (1951).

until it approached the limiting value for lattice diffusion at this temperature.

Table III shows clearly that for a given temperature, an increase in grain size lessens the contribution of boundary diffusion to the measured diffusion. (The ultimate effect would be achieved with a single crystal.) Conversely, for a given grain size, an increase in temperature achieves the same result; for the grain size used to obtain the results listed in Table II, a temperature slightly above 810°C would be necessary to reduce boundary effects to negligible values.

It should be made clear that the above statements concerning the data in Table III do not mean that the diffusion coefficients for boundary diffusion and lattice diffusion become equal at any point during grain growth at a given temperature. When grain growth proceeds to the extent that the boundary volume becomes quite small compared to the total volume, i.e., the available paths for boundary diffusion become few, then the boundary diffusion coefficient may be many times greater than the lattice diffusion coefficient without boundary diffusion contributing to any appreciable extent to the measured diffusion.

Little or no data exists with which to compare these reported results. Equation (7) agrees almost exactly with the results obtained by Buffington, Bakalar, and Cohen[15] on self-diffusion in γ iron, their equation being

$$D_{\gamma-\mathrm{Fe}}=0.58\exp(-67\,900/RT). \quad (8)$$

The reason for this close agreement is not clear. For the process of diffusion by a vacancy mechanism, the energy of activation may be considered to consist of contributions from the work necessary to form a vacancy plus the work necessary to move the diffusing atom from its equilibrium position in the lattice to the top of the potential barrier separating it from its next position. The magnitude of the activation energy would then depend on the equilibrium concentration of vacancies and the interatomic binding energies. Ordinarily it might be expected that in a solid solution local dislocations and perturbations of the lattice would exist, giving rise to an increased concentration of vacancies with a resultant lowering of the activation energy. In this case, however, the similarity between the effective metallic radii of iron, nickel, and chromium would tend to reduce this effect to a minimum. The close agreement between the activation energies in Eqs. (7) and (8) would therefore seem to indicate not only a similarity between the interatomic binding energies in pure γ iron and austenitic stainless steel, but also that the equilibrium concentration of vacancies in pure γ iron is not appreciably changed on going to a solid solution where γ iron acts as solvent, provided that the solute atoms produce no additional lattice strain.

Theoretical interpretations of D_0 are as yet incomplete[2]; it is at present considered to be a function of the atomic spacing, the vibrational frequency of the atoms, and the entropy change involved in formation of strains and vacancies in the lattice. For a complicated alloy such as 347 stainless steel, such information is not available; the reason for the identity of D_0 in Eqs. (7) and (8) is therefore not clear.

Garwick and Rosenquist[16] have carried out some investigations on the diffusion of iron in various steels. Their experimental technique, however, was somewhat questionable, consisting of sandwiching a lead foil between their radioactive iron and the steel, and allowing diffusion to proceed through the molten lead. In view of the extremely small penetration of iron into the steel at the temperatures used (600°C–970°C), the assumption that this setup in no way interfered with the diffusion process is rather doubtful. Their data do not permit any accurate calculations of Q or D_0, other than an order of magnitude estimation of Q of about 40 kcal/g atom. At 970°C, their measured D_{Fe} in 18–8 stainless steel was 2.9×10^{-11} cm^2/sec; according to Eq. (7) in this paper, D_{Fe} at this temperature should be 1.11×10^{-12} cm^2/sec. For a carbon steel containing 0.1 percent C by weight, they obtain at 970°C a D_{Fe} of 2.9×10^{-10} cm^2/sec; according to the data of Gruzin *et al.*,[17] who investigated the effect of small amounts of carbon on self-diffusion in iron, D_{Fe} in iron-carbon alloys containing 0.1 percent by weight of C is 1.1×10^{-12} cm^2/sec. In general, the work of Garwick and Rosenquist indicates that diffusion in austenitic steels is considerably slower than in cast iron or plain carbon steels. However, their results throughout show much faster diffusion rates than either the results reported in this paper for austenitic steels, or the results reported by Gruzin for low carbon steels.

ACKNOWLEDGMENTS

The authors wish to thank J. I. Hoover for his advice and suggestions during the course of the investigation. The assistance of E. O. Davis in several phases of the experimental work is also gratefully acknowledged.

[15] Buffington, Bakalar, and Cohen, *The Physics of Powder Metallurgy*, edited by W. E. Kingston (McGraw-Hill Book Company, Inc., New York, 1951), chap. 6, p. 103.

[16] J. V. Garwick and I. T. Rosenquist, Tidsskr. Kjemi, Bergvesen Met. No. 1, 6 (1951).

[17] Gruzin, Kornev, and Kurdyumov, Doklady Akad. Nauk S. S. S. R. **80**, 49 (1951).

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Measurement of Excess Ba in Practical Oxide Coated Cathodes*†

L. A. WOOTEN, GEORGE E. MOORE, AND W. G. GULDNER
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

(Received November 27, 1954)

These papers, designated I and II, discuss measurements of excess Ba (and/or Sr) in practical oxide coated cathodes. More than 300 cathodes of various designs were prepared, aged for periods up to 50 000 hours, tested thermionically, and then analyzed chemically.

Despite its great importance, the accurate measurement of excess Ba in practical cathodes is a problem whose complexity has not been adequately realized. Although the methods of the present experiments avoided many difficulties, limitations still exist. The difficulties and limitations, as well as the capabilities, can best be discussed in terms of typical measurements; we emphasize this aspect in I.

INTRODUCTION

USUALLY the physical and chemical state of a cathode is inferred from electrical measurements which allow great freedom for interpretation. Because excess Ba (and/or Sr) in or on the oxide cyrstals is usually considered of prime importance for thermionic emission, its direct measurement is fundamental for determining the state of the cathode. In all published[1–4] experiments, the cathode was exposed to a reagent believed to react only with excess Ba in the coating to produce an easily measured gaseous product as in

$$Ba+H_2O\rightarrow BaO+H_2. \quad (1)$$

To correlate excess Ba with emission, the analysis should measure, with a definite end point, all excess Ba or Sr present in the *cathode coat* during the final emission test and should be unaffected by Ba elsewhere in the tube or by other chemicals anywhere in the tube. Previous methods neglected important side reactions. For example, some authors[1,2] exposed the entire tube to the reagent, permitting reaction between the reagent and Ba or other reducing agents anywhere in the tube, but attributed the entire reaction to Ba in the cathode coat. Jenkins and Newton[4] avoided this error but pumped their tube and activated their cathode on the analytic system. They used the H_2 evolved in reaction (1) to measure Ba in the cathode. The H_2O was refrigerated in a trap adjacent to the tube. Much of the condensable gas given off by tube and cathode during exhaust could also condense, be given off again when the water was liberated for reaction (1), and cause unknown reactions before the excess water and other condensables were again frozen out.

EXPERIMENT

Method of Analysis for Excess Ba

Our analytical method employed reaction (1) because it occurs quantitatively at room temperature. Figure 1 shows the system used only for analysis. Some of its elements have been described.[5,6] The cathode was first exposed to water vapor and allowed to react. Surplus water was then frozen out; the remaining gas was then pumped into the calibrated pipette, and then circulated through CuO at 600°K and $Mg(ClO_4)_2$ at ambient in series. The H_2 first reduced CuO to form H_2O which was removed by $Mg(ClO_4)_2$. The PV product remaining was then measured in the calibrated pipette; the decrease is the H_2 removed. During circulation over CuO, CO was oxidized, without PV change, to CO_2 which was then removed by Ascarite. Typical results, shown below, indicate that substantially all gas evolved during exposure to H_2O was hydrogen.

Pipette sensitivity is limited by the same factors limiting McLeod gauges. A deflection of at least 1 mm is desirable; for a practical bulb and capillary this corresponds to about 10^{-5} mm pressure. In a volume of 300 cc, the minimum gas measurable was thus 0.003 cc mm. Since 1 cc mm of H_2 at 25°C corresponds to 7.4 micrograms of Ba, 0.022 μg of Ba was about the minimum measurable. For sensitivity in determining specific quantities such as excess Ba per unit cathode area, large cathodes were employed.

Total Emission Tests

"Total Emission" was given by the discharge, through the tube connected as a diode, of a capacitor charged to a voltage sufficient to draw about 0.9 amp/cm² from a sufficiently active cathode. The cathodes operated at practical temperatures (890 to 1090°K) during the test, which was so brief that the cathode was not damaged. The test adequately compares cathodes of a given type in tubes of the same

* Preliminary report, L. A. Wooten, Phys. Rev. **69**, 248 (1946).

† Figures, tables, and equations are numbered consecutively in this and the following paper.

[1] T. P. Berdennikowa, Physik. Z. Sowjetunion **2**, 77 (1932).

[2] H. Fritz, Mikrochemie **17**, 191 (1935).

[3] C. H. Prescott and J. Morrison, J. Am. Chem. Soc. **60**, 3047 (1938). These experiments deal with a very special cathode system.

[4] R. O. Jenkins, and R. H. C. Newton, J. Sci. Instr. **26**, 172 (1949).

[5] C. H. Prescott, and J. Morrison, Ind. Eng. Chem. (Anal. Ed.) **11**, 230 (1939).

[6] L. A. Wooten, Am. Soc. Testing Materials, Bull. **108**, 39 (1941).

TABLE I. Composition of metals used for filaments and cathode sleeves.[a] All compositions given in percent by weight.

	Filaments					Cathode sleeves		
Element	No. 1	Nickel No. 2	No. 3[b]	Pt[b]	Konal	Grade A Ni	Electrolytic Ni	Pt
Mg	0.09	0.12	0.005	...	0.006	0.049	<0.03	<0.005
Si	0.02	0.08	0.01	...	0.58	0.033	<0.03	<0.005
C	0.03	0.03	...	...	0.044	0.04		
Fe	0.24	0.21	0.01		6.61	0.082	0.03	<0.005
Mn	0.25	0.23	<0.001		0.08	0.096	<0.01	
Cu	0.14	0.11	0.005	<0.03	<0.03	0.015	0.18	<0.030
Co	0.05	0.05	0.001	...	17.48	~1.0	0.01–0.03	
Al	0.001	0.001	<0.001	...	<0.03	...	...	<0.005
Cr	0.01	0.01	0.001	...	...	...	...	
Ag	0.001	0.001	0.001	...	...	...	<0.03	0.005
Ni (by diff.)	99.17	99.16	99.97	...	73.2	98.62	99.62+	Pt 99.85+
Ti	...	...	...	...	2.45	0.054	...	...
S	...	...	...	...	0.03	0.0035	...	...
SiO_2				...			<0.005	
Zn				...			0.01–0.03	
Au				...				<0.03
Pd				...				<0.03
Rh				...				<0.03
Ca	...	...	...	...	...	...	...	<0.005
Na				...				<0.005

[a] Most analyses were by spectrograph. In many cases analysis was checked by other methods. A dash signified element was sought but not found. The sign < indicates that the element was observed but its abundance was less than indicated.
[b] Designated as pure Ni and pure Pt in text.

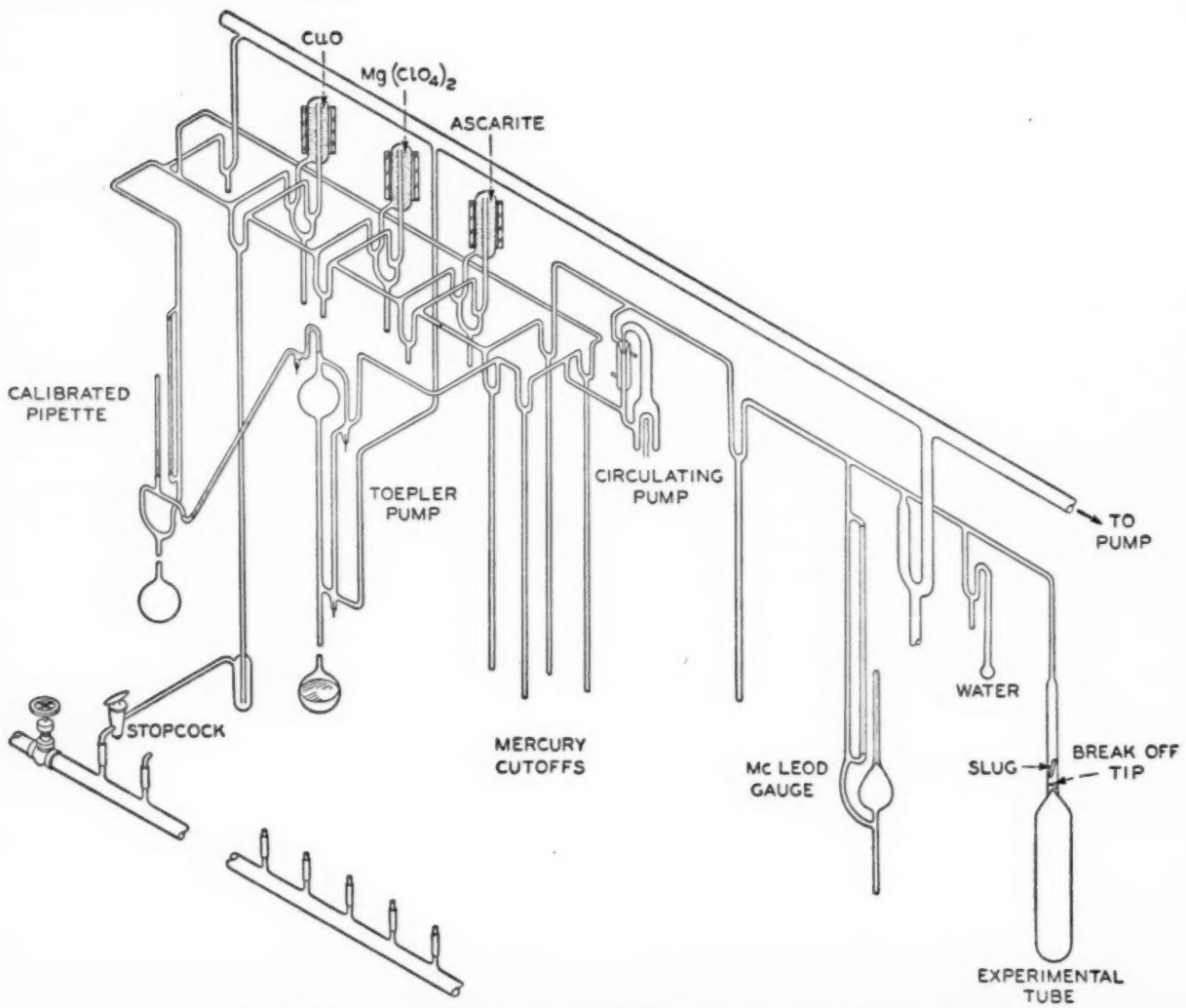


FIG. 1. Analytical system; pipette, toepler pump, and reagents discussed in references 5 and 6.

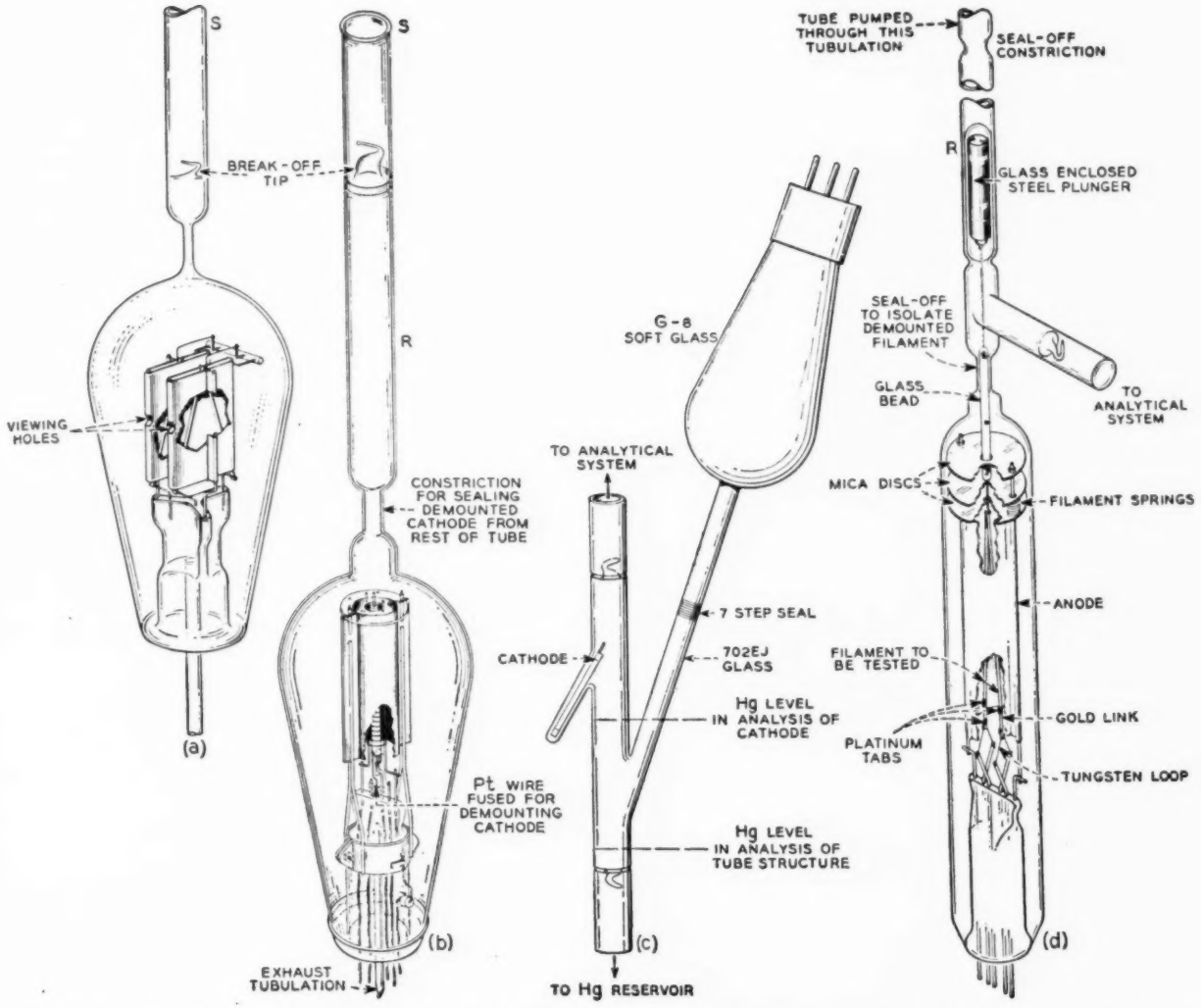


FIG. 2. Experimental tubes. (a) Filamentary cathode. Full wave rectifier. Entire tube exposed to H_2O. Coated filament area = 4.2 sq cm. (b) Demountable cathode triode. For analysis tube is inverted and 0.005-cm Pt wire burned out. Cathode then drops into analytical chamber R, which is then sealed from tube, and connected to analytical system at S. Cathode-diameter = 0.35 cm, length = 3.9 cm, coated area = 4.2 sq cm. Anode-bright or carbonized Ni, radius = 0.68 cm. Aged at 250 volts E_p and 0.020 to 0.030 amp I_p from cathodes of total emission above 200 ma. (c) Switch-back method. Cathode demounted and switched to indicated position. Entire structure then connected to analytical system and Hg reservoir. Bottom break-off tip then fractured and Hg raised to cut tube off from cathode. Top tip then fractured, and cathode tested with water vapor. Hg then lowered and tube exposed to water vapor. (d) Demountable filament tube. Filament width = 0.103 cm; thickness = 0.0038 cm; coated area = 5.3 sq cm; anode is carbonized Ni. Aged at 100 ma I_p. On completing emission tests, gold links (previously outgassed) were heated by tungsten loops and fractured by pull of solenoid on steel plunger.

geometry, but is less fundamental[7] than the term "total emission" might signify.

General

No getters were used. All cathode temperatures were determined by an optical pyrometer and corrected for spectral emissivity and bulb transmission.

For equipotential cathodes, unless noted, the spray coating contained $BaCO_3$ and $SrCO_3$ to give 35 mole percent BaO, and weighed 4 to 5 mg/cm² as oxides. For filaments the coat was applied by the drag method to the same density: the mixture contained 38.7 mole percent BaO. Table I gives the composition of the metallic supports.

As shown in Fig. 1, the experimental cathode was sealed to the analytical system which was conditioned[5,6] before cracking the break-off tip (Fig. 2).

Experimental Tubes

Figure 2 shows the four principal types.

RESULTS—PART I

Entire Tube Interior Exposed to H_2O

Figure 3 shows results for 6 tubes prepared as a homogeneous group as in Fig. 2a. One notes that (1)

[7] L. Nergaard, RCA Rev. 13, 478 (1952), discusses the difficulty in making a fundamental measurement of total emission from an oxide coated cathode.

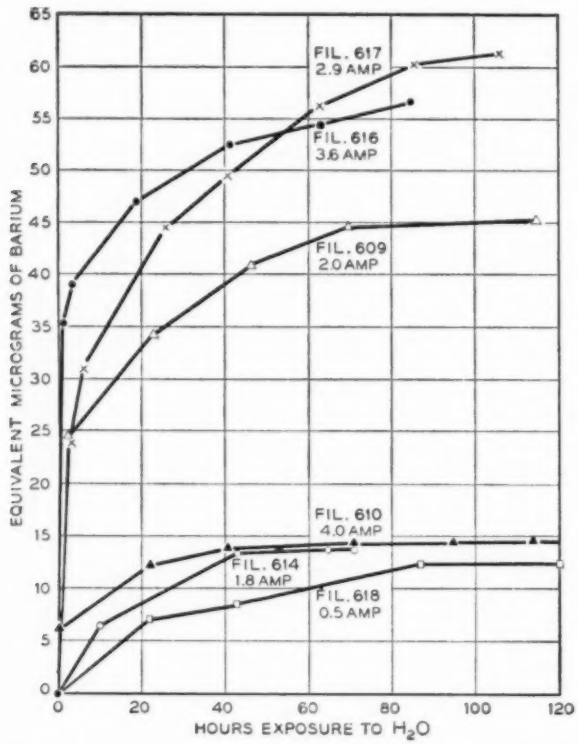


FIG. 3. Entire tube exposed to H_2O. Coated No. 1 Grade A Ni filaments. Aged 5 minutes and then operated at 370°K during exposure to H_2O. Emission indicated below each filament number.

reaction with H_2O is slow without a definite end point; (2) the initial rate varies greatly from tube to tube; and (3) thermionic emission does not correlate with the hydrogen produced during any exposure time; thus tube 610, with the highest emission, was a poor fourth in H_2 produced. Features (1) and (2) contrast strongly with accurate measurements[8,9] made on the total amount of alkaline earth oxide reduced by chemically active filament support metals. Behavior during such measurements is illustrated in Fig. 4 involving twelve tubes, as in Fig. 2a, except that each contained *only* the coated filament. The H_2 liberated is plotted for each tube against time of exposure to H_2O on a logarithmic scale. The end point was definite within 15 minutes in all but one case‡ even for 362 micrograms of Ba corresponding to 49 cc mm of H_2. These results show that when Ba is readily available to water vapor, reaction at ambient is conveniently rapid and the hydrogen evolved reliably measures available excess Ba. Spectrochemical tests showed that essentially all the Ba here had deposited on the bulb.

Many detailed tests showed that reactions other than (1) caused the irregular behavior of Fig. 3. We next discuss one such test.

Illustrative Test of Side Reactions

Tubes with anodes were prepared as in Fig. 2a. Some contained a coated and others an uncoated filament. During exhaust, the anodes were heated to 1125°K for 6 minutes and the filaments to 1300°K for the final 3 minutes.

Figure 5 compares two analyses. At least 60 hours exposure to water at ambient were required for an end point. Comparable amounts of H_2 evolved up to 139 hours in the two tubes, yet there was no Ba whatsoever in tube 485. Then, *while exposed* to H_2O, each filament was heated at 350°K for 21 hours. This produced slight additional H_2. The filaments were then heated for 85 hours to 500°K while exposed to H_2O. For the uncoated filament, this gave more H_2 than the entire previous exposure, but gave only slight additional H_2 for the coated filament.

After 250 hours exposure, the total gas evolved was passed over the hot CuO and $Mg(ClO_4)_2$; substantially all disappeared, indicating H_2. Obviously, reactions other than (1) produced all the H_2 of Fig. 5 in tube 485. Two important side reactions are thermodynamically possible: (a) Impurities commonly present in Ni, such as Mg, Ti, or Si can[10,11] liberate H_2 from water at ambient. (b) Ni itself can[12] liberate H_2.

Many detailed tests show that reactions of type (a) occur readily at ambient with anode or cathode; the composition of the anode was always essentially that of No. 1 Ni. Similarly, detailed tests show that Ni itself does not react perceptibly from ambient up to more than 100°C, although reaction at ambient is thermodynamically possible.[12] Uncoated pure Ni always reacts at 200°C but Fig. 5 and other tests show that reactions of type (a) or (b) are inhibited by the coating and processing of practical cathodes.

DEMOUNTED CATHODES AND FILAMENTS

Figure 6 illustrates typical analyses of demounted cathodes from tubes as in Fig. 2b. The H_2 evolved is much less than in preceding examples. The uppermost curve is for a coated Grade A Ni cathode. Note the end point after 16 hours at ambient, and that 30 minutes at 200°C caused further reaction in three different tests.

[8] G. E. Moore, J. Chem. Phys. **9**, 327 (1941).
[9] Moore, Allison, and Morrison, J. Chem. Phys. **18**, 1579 (1950).
‡ Occasionally while filaments were heated, the bulb was surrounded by the induction coil, which carried the same current as used for heating the anodes during commercial exhaust. Because no anodes were used, the field of the coil did not heat the filament perceptibly, but more Ba was always produced. This is illustrated by comparing the sample, shown by solid points in Fig. 4b, with the two experiments shown in the bottom curve. About 20 times as much Ba was produced but the end point was indefinite. "Electrodeless" excitation of gas may have been responsible. The effect was not studied beyond ascertaining its existence in several tubes.

[10] F. D. Rossini, "Selected Values of Chemical Thermodyanmic Properties," National Bureau of Standards (1950–1951).
[11] A. H. White, J. Appl. Phys. **20**, 856 (1949).
[12] R. N. Pease and R. S. Cook, J. Am. Chem. Soc. **48**, 1199 (1926).

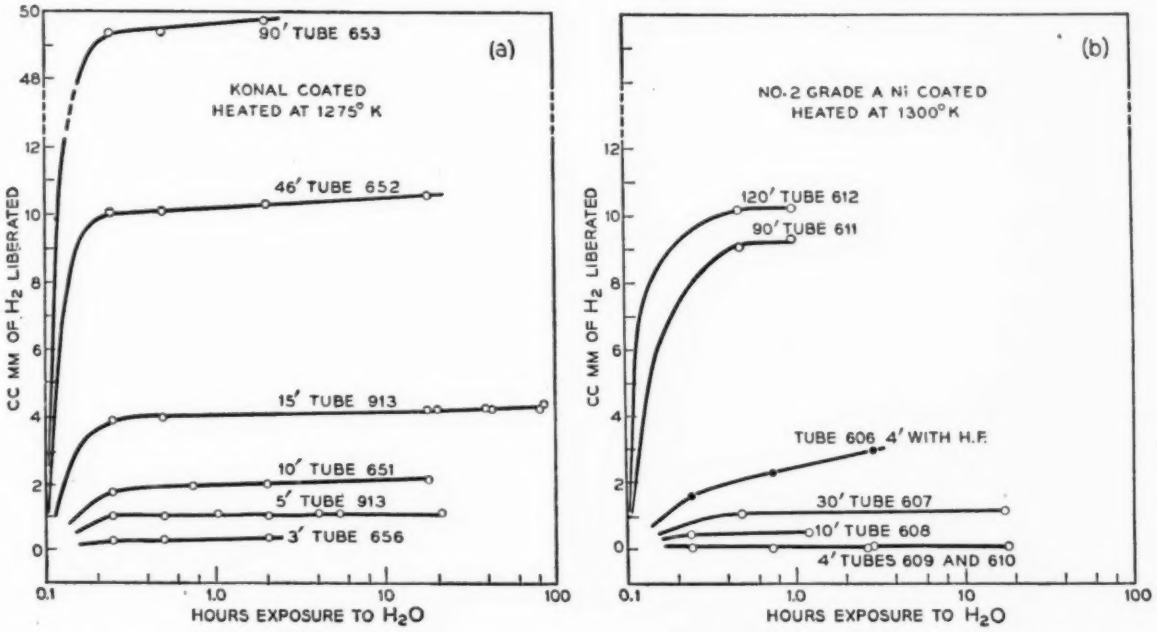


FIG. 4. Indication that method will measure correctly readily available Ba, etc. Part (a) filament support is Konal, a Ni-cobalt-Fe alloy, rich in Ti and Si (see Table I). Note break in ordinate scale. Part (b) filament support is No. 2 Grade A Ni. Each filament heated on exhaust system for time and temperature indicated before connection to analytic system.

Similar tests were made on 19 "blanks" of uncoated electrolytic Ni, or Pt, cathode sleeves prepared in tubes as in Fig. 2b except that anode, grid, etc., were omitted. Figure 6 gives data on three. EK 121 of Pt was treated at 925°C in H_2 for 15 minutes before assembly, then at 1000°C for 5 minutes during exhaust, and finally aged at 675°C for 1584 hours. It produced much less H_2 at ambient than the coated cathode DC 397. It was then heated at 200°C for 5 minutes and re-exposed to H_2O, reacting comparably to the coated cathode. Only one blank gave no observable reaction at either ambient or at 200°C. Two additional blanks are shown, one using electrolytic Ni and the other a platinum sleeve previously baked at 1300°C in air for 2 hours before assembly and then aged at 700°C for 15 minutes after exhaust. The electrolytic Ni gave a low reaction at ambient but rapid reaction (not shown) at 200°C.

Figure 6 illustrates three features common to experiments with demounted cathodes: (a) Although reduction by the blanks was usually considerably less than by corresponding coated cathodes, we were unable to prepare at will Ni or Pt blanks which gave no measurable reduction of H_2O. (b) The H_2 evolved by the coated cathode is far smaller than for the entire tube.

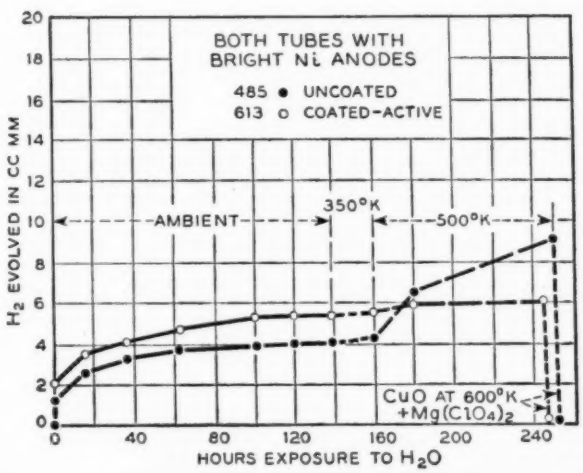


FIG. 5. Comparison of coated and uncoated No. 1 nickel filaments in bright Ni anodes. Emission of coated filament at 1020°K, after aging 16 hours, was 3.60 amperes = 870 ma/cm².

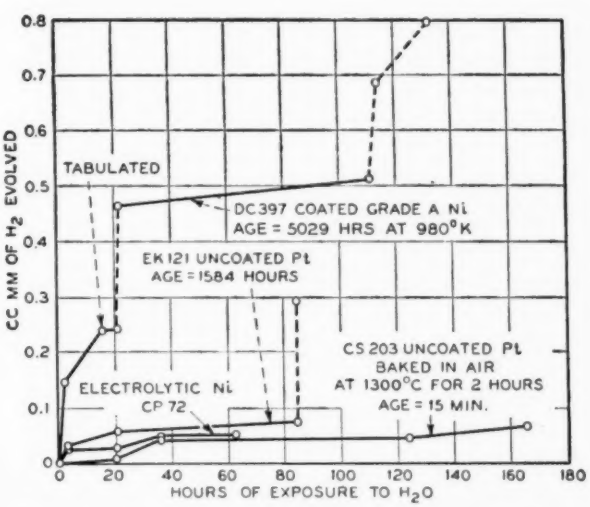


FIG. 6. Hydrogen evolution by demounted cathodes. Exposure to H_2O was at ambient, but in the sections shown dotted the cathode had been heated 30 minutes at 200°C in vacuum before exposure to water vapor.

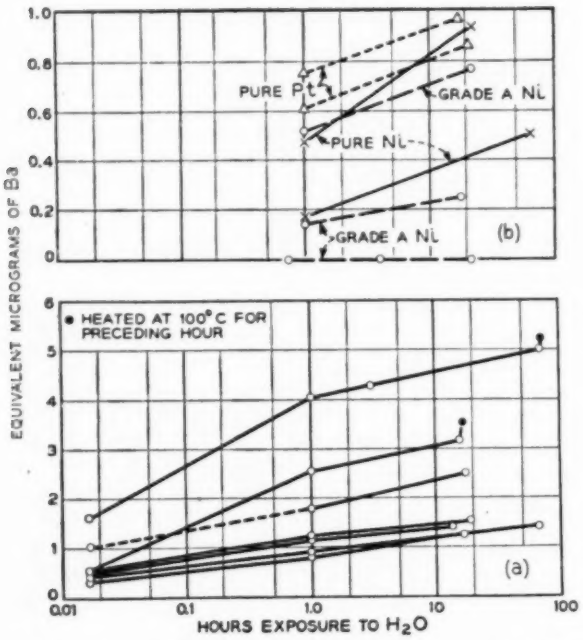


Fig. 7. Comparative behavior of various demounted cathode systems in water vapor tests. Some curves are drawn dotted for identification only. Note that for the demounted filaments of Fig. 7b, the Ba in the coatings on Pt supports was larger than for Ni. The Pt was extraordinarily pure (see Table I).

(c) Heating to 200°C or above invariably caused much additional evolution of H_2 especially from uncoated cathodes.

Comparison of Coated Cathodes and Filaments

Figure 7(a) illustrates the behavior of coated electrolytic Ni sleeves. Their ages ranged from 15 minutes to 1268 hours. Two features are noteworthy: (a) Measured Ba varies almost linearly with log of time, without an end point in these coordinates. The H_2 usually increased about threefold for a thousand fold increase of time. However, when plotted on a linear time scale, the rate for a demounted cathode seems to reach an end point This is illustrated by the first three points in Fig. 6 for DC 397. The point marked "Tabulated" and the two adjacent were linear on a plot of H_2 against log of time. (b) The solid dots indicate demounted cathodes which were heated to 100°C by maintaining boiling water around the glass tubing, during the hour of exposure. This never increased the hydrogen evolution appreciably.

Figure 7b deals with demounted filaments. The nickel or platinum was completely covered with the oxide coating, and the mass ratio of coating to support metal was much larger than with equipotential cathodes. The platinum had essentially no impurity or interface, so that the reaction must be with Ba in the oxide. Note that one filament, although active thermionically, indicated no Ba.

For correlation with emission, the H_2 from 16 hours exposure was taken as the measure of Ba content. There are dangers in overexposure, such as slow reaction with the Ni itself or with impurities in the Ni. Reasons for the decision are: (a) Readily available Ba, Sr, or Mg produced in chemical reactions, as in Fig. 4, react quantitatively with H_2O at ambient within an hour. (b) By present theories[7,13] Ba significant for thermionic emission, lies close to the outer surface of the oxide crystals and should be nearly as accessible to the H_2O as in the experiments of Fig. 4. (c) Exposure to H_2O at ambient for 16 hours changes the (BaSr)O phase to the hydroxide, as shown by conversion of the coating to a mushy phase and by[4] x-ray diffraction.

DISCUSSION

Ba_2O is sometimes considered in studying properties of BaO, but evidence indicates[14,15] it does not occur. Reaction (1) would not distinguish a molecule of Ba_2O from an atom of Ba. Evolution of one molecule of H_2 could indicate either an atom of excess Ba or an oxygen vacancy. These differ only in diffusion mechanism.

CONCLUSIONS

Because of side reactions, the accurate measurement of Ba in a cathode coating is more difficult than previously indicated. Interference by impurity in the Ni or Pt support metal is troublesome. Ni itself does not react with H_2O below about 150°C. With a coated cathode, reaction (1) does not give an absolute end point, but the H_2 formed on 16 hours exposure is taken as the measure of Ba content significant for emission. This H_2 is probably essentially all formed by Ba in the coating, as judged by comparative tests of cathodes and filaments.

---

[13] Conyers Herring (personal communication); also internal report issued in 1946.

[14] M. Schriel, Z. anorg. Chem. **231**, 313 (1947).

[15] Sproull, Bever, and Libowitz, Phys. Rev. **92**, 77 (1953). This gives references to previously published work by the Cornell group.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Excess Ba Content of Practical Oxide Coated Cathodes and Thermionic Emission *

GEORGE E. MOORE, L. A. WOOTEN, AND J. MORRISON
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey
(Received November 27, 1954)

Using methods given in paper I, emission and Ba content are given on (a) 125 coated cathodes, (b) 15 coated filaments; H_2 evolution by 19 uncoated cathodes was also measured. For analysis, each was isolated from the remainder of its tube. Coatings varied in composition from pure SrO to pure BaO and in density from 1 to 5 mg per cm². Ba was also measured in cathode coatings removed from their support metal. Essentially all cathodes contained excess Ba, but no cathode structure gave any correlation between Ba content and emission.

These experiments contribute the following toward understanding practical oxide cathodes: If excess Ba (or Sr) is important, the significant amounts are probably less than one atom of Ba in 10^6 molecules of (BaSr)O. Because essentially all cathodes, including some with extremely low emission, contained at least this amount, we believe that performance is seldom limited by excess Ba. The concentration of Ba (or Sr) was not affected by using pure BaO or SrO coatings or Pt supports, although each of these steps led to very low emission; the highest emissions were obtained by operating at relatively low temperature during life, but this retained no more Ba in the coat. These major effects require interpretation.

INTRODUCTION

THE purposes of this work were the following: (1) to measure in various cathode structures, the excess Ba content of the coating in each of several samples; (2) to measure the thermionic emission of each sample; (3) in each cathode species, to attempt a detailed correlation between Ba content and emission; (4) to investigate the influence on Ba content and emission of (a) age of cathode, (b) anode material, (c) operating temperature, (d) variations in cathode coating, (e) composition of cathode support metal, and (f) treatment of support metal before coating.

EXPERIMENT

Exhaust processes were performed on "trolley" systems; the tubes were first baked at 425°C for 3 hours. The anodes were then heated to 1125°K for 6 minutes by induction; during the final 3-minute interval, the cathodes were heated to 1300°K.

The cathodes were activated, aged for periods up to 50 000 hours and tested for emission at intervals. Ba content was determined by reaction (1). Tube structures are indicated in Fig. 2. Thermionic quality was never in statistical control in the sense of a satisfactory commercial product, probably because of the omission of all getters. The best cathodes were of high commercial quality, but many other cathodes would have been useless in any practical application. This wide range in emission is advantageous for correlation.

RESULTS—PART I

Practical Cathodes and Filaments and Variations

(*1*) *Demountable Grade A Ni Equipotential Cathodes.*—Table II, Part A, represents tubes with carbonized Ni anodes and Grade A Ni cathodes. Samples are tabulated according to age, listed in the second column. Total emission at 980°K is given in the third column. Experience indicates that 800 ma forecasts a life exceeding 10^4 hours. The fourth column gives the Ba equivalent of the H_2 evolved; note the relatively low abundance in CC 493 with the highest emission (1780 ma). For comparison, Part B represents cathodes from the same lot, in tubes having bright nickel anodes; there is no significant difference in mean results for the two groups. For each group, the mean and rms deviation are listed.

The conclusion of Jenkins and Newton[4] that the cathode retained only about one-eighth of the total Ba originally liberated in the coating was extended by tests measuring the Ba in the cathode and in the remainder of the tube and the ratio to that in the cathode, respectively. This ratio ranged from 271 to 552 which is surprisingly uniform, considering that ages range from 0.5 to 56 000 hours. Table II, Part C, gives these results.

All equipotential cathodes have a coated area of 4.2 cm² and a nominal weight of coating 4 to 5 mg per cm² as carbonates.

For most subsequent groups, the data are merely summarized.† Usually included will be only the number of cathodes, range of age, mean and rms total emission, together with mean and rms Ba content. Detailed data were studied for all groups by E. L. Kaplan of this laboratory. He found no significant correlation in any group between Ba content and emission.

(*2*) *Demountable Cathodes—Variation from Standard Practice.*—Table III includes several variations from the approximately commercial techniques of Table II, again using the structure of Fig. 2(b). In Part A the cathodes were aged at 890° and in Part B at 1060°K, 90° cooler and 80° hotter, respectively, than for the cathodes in Table I. The emission of Group A, measured at 890°K, was multiplied by 3.84, the mean ratio of emission at 980°K to that at 890°K for cathodes of this type. Their

* The figures, tables, and references in this and the preceding paper are numbered consecutively.

† Detailed tables will be gladly sent to anyone interested.

TABLE II. Coated grade A nickel cathodes.

Tube No.	Age hr	T.E. ma	Equiv. Ba in micrograms on 16-hrs, exposure to H_2O
	Part A. Carbonized Ni anodes.		
CC 309	0	1700	0.65
CC 492	0	1180	0.35
CC 493	0	1780	0.45
CM 90	1/4	1060	0.48
CM 103	1/4	200	1.44
CP 1	1/4	1330	1.18
CP 7	1/4	265	0.44
DC 755	1/4	410	0.22
CC 489	20	1700	0.29
CC 490	20	1000	0.40
DC 756	562	850	0.44
CP 11	907	610	0.67
CP 12	907	1270	0.38
CM 97	1048	1260	0.33
CC 488	1268	1480	0.34
CC 491	1268	1180	1.01
CP 2	1304	252	0.44
CP 3	1963	1290	1.98
CM 101	2354	1200	1.11
DC 396	2926	1020	0.95
DC 394	2968	1320	2.14
DC 401	2968	1300	2.60
DC 397	5029	1360	1.07
	Mean	1088	0.84
	rms	461	0.66
	Part B. Bright Ni anodes.		
CM 70	1/4	1120	0.47
CC 312	20	900	0.52
CC 313	20	1560	0.50
CM 61	66	1550	0.70
CM 62	1048	1280	0.61
CM 71	1048	1480	1.32
CM 69	5959	730	3.20
	Mean	1231	1.05
	rms	328	1.00

Part C. Switch-back technique, all anodes carbonized.

			Cath.	Tube	Ratio
DC 753	1/2	150	0.23	110.	472
DC 756	62	850	0.47	207.	446
DE 359	1655	1310	0.43	213.	493
DE 362	1655	1160	0.32	86.0	271
DE 361	56532	884	0.67	366.	552
	Ave	871	0.42		447
	rms	446	0.17		42

specific emission was the highest obtained but their Ba content does not differ significantly. The cathodes aged at 1060°K were tested directly for total emission at 980°K. Although their mean emission was only 37 percent that for the cathodes of Part A, their Ba content was greater, but probably not significantly so.

The cathodes of Part C had only 1 mg of coating per cm^2, about $\frac{1}{4}$ to $\frac{1}{5}$ the surface density for other groups.

TABLE III. Variations from usual temperature or coating.

Part	Variation	No. of tubes	Age	T.E. Mean	T.E. rms	Ba μg Mean	Ba μg rms
A	890°K	4	2289–5350	3230	876	1.27	0.97
B	1060°K	4	2926–3091	1213	90	1.42	0.93
C	1 mg/cm²	5	1/4–7896	693	113	1.50	0.82
D	Pure BaO	3	90–2715	81	86	1.90	0.73
E	Pure SrO	3	884	278	110	1.03	0.49

TABLE IV. Electrolytic Ni cathode sleeves.

Part	Variation	No. of tubes	Age	T.E. Mean	T.E. rms	Ba μg Mean	Ba μg rms
A	Dry H_2	8	0–1268	1323	339	2.13	1.19
B	Wet H_2	10	0–52840	821	322	1.23	0.64

Although emission was somewhat lower than in preceding groups, the Ba was as abundant, indicating a mean concentration 4 to 5 times the mean for preceding groups.

Parts D and E give results for cathodes coated only with BaO, or with SrO, respectively. The amount of excess Ba or Sr was equivalent to that in cathodes coated with the usual mixture, but their total emission was much lower, especially with BaO.

(*3*) *Equipotential Cathodes Using Electrolytic Nickel Sleeves.*—Table IV deals with cathodes of electrolytic nickel. In Group A, the cathode sleeves were treated for 15 minutes in dry H_2 before coating, and the cathodes of Group B were treated in *wet* H_2 in order to "fix" any reducing impurities in the nickel as oxides. The purpose was to decrease the Ba content of the cathode coat and the results indicate both a lower Ba content and lower emission by amounts that might be significant. However, one cannot claim correlation because (a) The tubes of Group B were assembled and processed about one year later than those of Part A. Despite the effort to utilize homogeneous parts, small variations certainly entered. (b) There is again no correlation between emission and Ba content among individual tubes of either group.

In any repetition of the present experiments, closer comparisons of such features of preparation should be made.

(*4*) *Demountable Cathodes Using Platinum Sleeves.*—Table V deals with Pt cathodes aged at three temperatures. The Ba content of their coating is significantly higher than in most of the lots having Ni sleeves, but their emission is far lower.

(*5*) *Blanks for Equipotential Cathodes.*—To evaluate H_2 produced by the sleeve itself, 19 "blanks" were tested.

TABLE V. Platinum cathode sleeves.

Tube No.	Hr	T.E. ma	Ba μg	Cath. temp.
DC 65	0	329	0.59	980
DC 67	0	400	0.96	980
DC 236	1304	82	2.14	980
DC 235	1500	210	1.41	980
DC 66	2937	11	5.16	980
DC 72	2937	91	2.95	980
DL 290	52840	5	2.79	980
DC 242	1304	34.2	1.60	890
DC 239	2937	5.3	2.06	890
DC 240	13614	31.6	1.65	890
DC 71	2934	200	1.70	1060
	Mean	127	2.09	
	rms	138	1.23	

Results are shown in Table VI. Each tube [Fig. 2(b)] contained only the uncoated cathode with its heater. Because H_2O reacts with impurities in Grade A Ni, blanks were measured only for platinum and electrolytic Ni. There is a clear trend toward lower "equivalent Ba" in cathodes treated in oxidizing atmospheres. We believe that the electrolytic nickel blanks give an upper limit to the H_2 formed by the support metal in coated electrolytic, or Grade A, cathodes, and similarly for Pt. No correction for blanks was made.

(*6*) *Demounted Filaments.*—Two possibilities for error exist in tests of equipotential cathodes: (a) Heater—insulator reaction in which the cathode was heated internally by tungsten in contact with alumina or beryllia insulation. The tungsten, operating at approximately 2000°K, reacts with the insulation to liberate Al or Be, as in analogous reactions.[8,9] These uncontrolled amounts of Al or Be then deposit during life on the inner surface of the sleeve and may produce H_2 in addition to that from Ba in the cathode coat. Absence of a general trend toward higher apparent Ba content with age argues against heater reaction causing serious error. (b) The interior surface of the sleeve and $\frac{1}{8}$ in. at each end of the exterior were uncoated. Reduction of H_2O by impurities which diffuse to the uncoated surfaces could introduce error. We believe this is small because CO_2, etc., during exhaust of coated cathodes would oxidize impurities to an extent greater than is observed with uncoated sleeves.

TABLE VI. "Blank" measurements on uncoated sleeves.

Part	Variation	No. of tubes	Age	μg equivalent Ba at 16 hr Mean	rms
A	Electrolytic Ni 925°C in dry H_2 for 15 min	6	1/4–1584	0.297	0.145
B	Platinum 925°C in dry H_2 for 15 min	9	1/4–50 000	0.572	0.645
C	Platinum 1300°C in Air for 2 hr	4	1/4	0.15	0.17

The tube of Fig. 2(d) should eliminate these possibilities. The filament is entirely coated and directly heated; results are shown in Table VII. The mean Ba content for three different filament support metals is less than for the equipotential cathodes but again no correlation exists between emission and Ba content. For EN 178, no Ba was observed but its emission seemed unimpaired. Filaments of Part C employed Pt of extraordinary purity; although their mean Ba content was higher, their emission was significantly lower than for the nickel filaments.

RESULTS—PART II

Tests in Which the Coating is Removed

Because unknown factors associated with the support metal may introduce error in determining Ba in the coat, experiments by two methods were performed in which the (BaSr)O coating was removed in vacuum and then subjected to H_2O.

TABLE VII. Demounted filaments.

Tube No.	Age hr		T.E. ma	μg equiv. Ba
Part A. Grade A nickel support metal.				
EA 403	0		1000	0.26
CH 647	16		850	0.72
HH 45	17		542	0.75
HH 46	18		784	3.47
EP 102	24		1150	0.17
DH 648	27		900	1.25
EN 178	95		1230	0.00
HL 1400	128		100	0.10
HM 1082	814		1878	0.75
HH 44	841		1250	0.42
HP 518	888		3565	0.24
		Mean	1205	0.74
		rms	902	0.98
Part B. Pure Ni supports.				
HL 1410	128		1925	0.90
HM 923	215		1430	0.34
		Mean	1678	0.62
Part C. Pure platinum supports.				
KD 11	120		70	0.80
KD 12	912		730	0.95
		Mean	400	0.875

Konal Filaments in Structure of Fig. 2(a)

After considerable aging, the (BaSr)O will sometimes drop off Konal; this was accomplished in two cases. The coating was shifted to the analytical chamber which was then sealed from the remainder of the tube and subjected to H_2O, with results of Table VIII. The Ba found was typical for measurements on the entire filament or cathode. Emission phenomena are discussed in Appendix A.

TABLE VIII. Coating only analyzed.

Part A. Coating dropped off Konal cathodes.

Tube No.	Age hr	Emiss.	μg equiv. Ba	Emiss./ cm²	Total coat collected	Wt percent of Ba
DD 344	185	108.0	3.04	26.7	16.0 mg	0.019
PE 185	40	2 ma	0.229 μg	0.6	13.2	0.0017

Part B. Coating scraped off. Emission at 3000 volts

Tube No.	Age hr	Emiss./ cm²	Total coat collected	μg equiv. Ba	Wt percent Ba
VD 156	7½	2.08 ma	20.52 mg	0.96	0.00469
VC 98	20	63.8	6.84	1.53	0.0224
VD 157	24	100.0	19.14	0.73	0.0038
VB 243	48	26.7	13.02	2.29	0.0176
VC 192	48	61.0	6.96	1.54	0.0221
UM 31	120	46.7	2.44	0.47	0.0193
				Mg 1.25	0.0149
				rms 0.66	0.0078

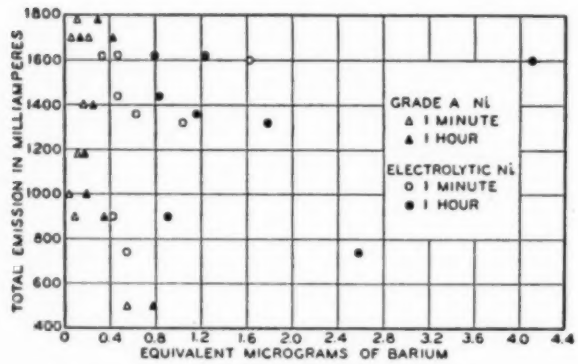


FIG. 8. Correlation chart for emission *vs* Ba content measured on one-minute and one-hour exposure to H_2O.

Experiments in Which the Coating is Scraped Off Support

These employed a tube with large coated surfaces of sheet Ni. The coating was dislodged by scraping with a rod which entered the tube through a sylphon bellows. The coating removed was shaken into the analytical chamber which was then isolated by crimping a section of copper tubing, eliminating the need for fusing glass. Emissions were measured by 3000 volt pulses at 1090°K. Results, in Table VIII (B), show that the specific Ba content of the cathode agreed with previous tests.

RESULTS—PART III

Exposure to H_2O for Less than 16 Hours

In present theories[7,13], Ba significant for thermionic emission lies close to the surface of the oxide crystals and might be especially accessible to the H_2O. Absence of correlation between emission and Ba measured for 16 hours, suggested comparing the H_2 evolved after one minute and after one hour with total emission. Figure 8 shows the results for 8 cathodes with Grade A Ni and 8 with electrolytic Ni sleeves. (1) There was no correlation between thermionic emission and Ba content in either lot measured after either one minute or one hour. (2) After either one minute or one hour, the electrolytic nickel cathodes showed significantly more Ba than the Grade A, without significant difference in emission.

DISCUSSION

Statistical Treatment of Data

Most of Kaplan's statistical analysis utilized 10 groups of 5 or more cathodes each, comprising a total of 87. Table IX and the following summarize his conclusions.

The strongest correlation between Ba content and emission was *negative* and occurred for the 7 cathodes in Pt sleeves of Table V, aged at 980°K. The coefficient[16] for this group was −0.82: if this had been the only group, the chance for such a result would be only 2 percent, but for 10 groups, the chance in one group is 20 percent. Six of the 10 groups showed small positive correlations. Considering the Ba content in all groups, regardless of age, there is virtually no evidence of correlation between emission and Ba content.

To investigate the effect of age on both emission and Ba content, each of 7 groups was divided into two parts according to age; mean emission and Ba contents were compared by the *t*-test.[16] No significant change with emission was found, but in some groups Ba content definitely tended to increase with age. This was most marked for two cathodes aged 1268 hours in Table IV(A). Their mean Ba content exceeded that of the other 6 cathodes by 2.29 micrograms, which would occur by chance among 7 comparisons only about 2 percent of the time. Tables VI(A) and VIII(B) show similar effects which are less definite and therefore not statistically significant. Two of the 7 groups show small *decreases* in Ba content with increasing age; among these is Table VII for the filaments. In equipotential cathodes any error caused by evolution of Al or Be from chemical reaction between the tungsten heater

TABLE IX. Statistical examination of data.

Group	Age	n	Ba Mean	Emission *vs* Ba b	A_b	r	Significance* Ba *vs* age (percent)
III A	All	23		90	430	0.11	
	≤ 20	10	0.58				10
	≥ 562	13	1.03				
III B	All	7		−280	430	−0.55	
	≤ 66	4	0.55				15
	≥1048	3	1.71				
III CC	All	5		1030	8300	0.39	
IV C	All	5		50	430	0.37	
V A	All	8		−20	430	−0.07	
	0, 20	6	1.55				0.3
	1268	2	3.84				
V B	All	10		90	580	0.18	
	≤2326	5	1.48				25
	≥4387	5	0.98				
VI A	All	7		−80	100	−0.82	
	≤1500	4	1.28				3
	≥2937	3	3.63				
VI D	All	5		0.002	0.008	0.59	
VII A	All	11		−170	980	−0.19	
	≤ 27	6	1.10				20
	≥ 95	5	0.30				
VIII B	All	6		0	270	0.03	
	≤ 24	3	2.06				20
	≥ 48	3	3.93				

b = slope of least-squares straight line for estimating emission from measured Ba content; 99 percent probability (per group) that true slope lies in $b \pm A_b$. The fact that A_b is in all cases greater than b means that no correlation exists.
r is correlation coefficient.
* Significance—the probability of being incorrect in stating that a correlation exists between Ba and age in the given grasp, without regard to sign.

[16] See, for example, Croxton and Cowden, *Applied General Statistics* (Prentice-Hall Inc., New York, 1941).

and its insulation would increase with age and might cause an apparent trend of Ba content with age in equipotential cathodes, but not in filaments. However, the effects are not sufficiently marked to permit a definite conclusion.

Since Ba content may increase with age, the data were examined to see if higher emission were indeed related to higher Ba content, but this effect was counteracted by some factor associated with age. For this, 41 pairs of cathodes (in 14 groups) were so selected that the two cathodes of any pair belonged to the same table and were of the same age, or very nearly so. Twenty pairs showed positive correlation between emission and Ba content, while 21 pairs gave negative correlation.

From these tests, Kaplan concludes "There is no evidence of an over-all effect of Ba (as measured in these experiments) on emission even at constant age." His examination is partially summarized in Table IX.

Median Ba Content

Several authors have published theoretical or experimental values of donor concentration in oxide cathodes, usually implying the donors to be Ba or Sr. The present work suggests a "probable" concentration of Ba. Because distribution curves of Ba content "tail off" toward higher values, and because residual errors of method tend to declare too much Ba, the *median* concentration is probably more significant than the mean. Figure 9(a) shows the distribution of 56 cathodes with Ni sleeves and includes all samples in Parts A and B of Table II and all samples of Tables III and IV. The median concentration is 6 atoms of Ba in 10^5 oxide molecules. The mean "blank" for Ni cathode sleeves would decrease this by about 25 percent if the entire blank applied. Probably more representative is Fig. 9(b) for the 15 filaments of Table VIII. The median for this distribution gives 2 atoms of Ba for 10^5 molecules of oxide.[17]

Sproull, Bever, and Libowitz[15] found 0.03 to 0.04 percent excess Ba in single crystals of BaO, colored blue by exposure to Ba vapor at 1100°C, considerably more than our means but no larger than several individual samples. Although Jenkins and Newton[4] state concentrations not far from ours, our rms spreads were much larger than theirs. This may be the result of extraneous gas in their method.

Significance for Theory

Although excess Ba in usually considered necessary for high emission, no satisfactory experimental evidence has ever been published. The present program was undertaken to provide such evidence by correlating

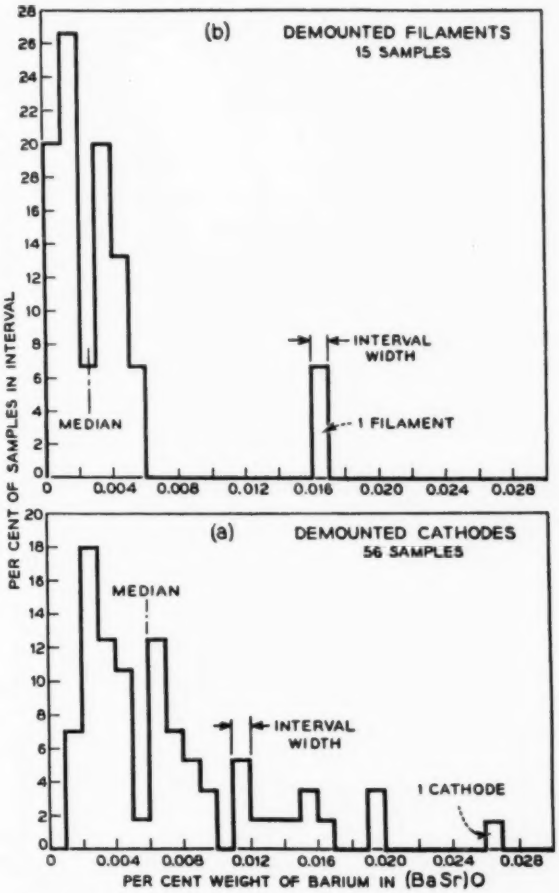


FIG. 9. Distribution chart of specific Ba content: (a) demounted equipotential cathodes and (b) demounted filaments.

emission and Ba content. Failure to observe correlation does not disprove the theory; however, the measurements have consequences such as: (1) if excess Ba is of major importance for thermionic performance of practical cathodes, correlation probably lies below our sensitivity, which was about 0.005 $\mu g/cm^2$ or about 10^{-4} percent by weight (or one atom of Ba in 10^6 molecules of oxide); (2) the poor emission of many cathodes with relatively high Ba content strongly indicates that in practical cathodes, factors other than Ba content (such as cathode temperature) are decisive; (3) theoretical work is needed to understand such results as the poor emission, despite relatively large quantities of excess Ba from coatings on platinum supports and of pure SrO or BaO.

ACKNOWLEDGMENTS

The writers acknowledge gratefully the help received from many others, especially from H. W. Allison, E. K. Jaycox, and A. E. Ruehle, who performed some of the experimental work; from W. D. Stratton and J. W. West, who supervised the construction of experimental tubes; and from K. G. McKay and E. L. Kaplan, with

[17] Compare with values published by Eisenstein (*Advances in Electronics*, Academic Press Inc., New York, 1948) Vol. I. The quotation that $n_b = 10^{18}/cm^3$ pertains to our work, but is too high, and is probably a misprint.

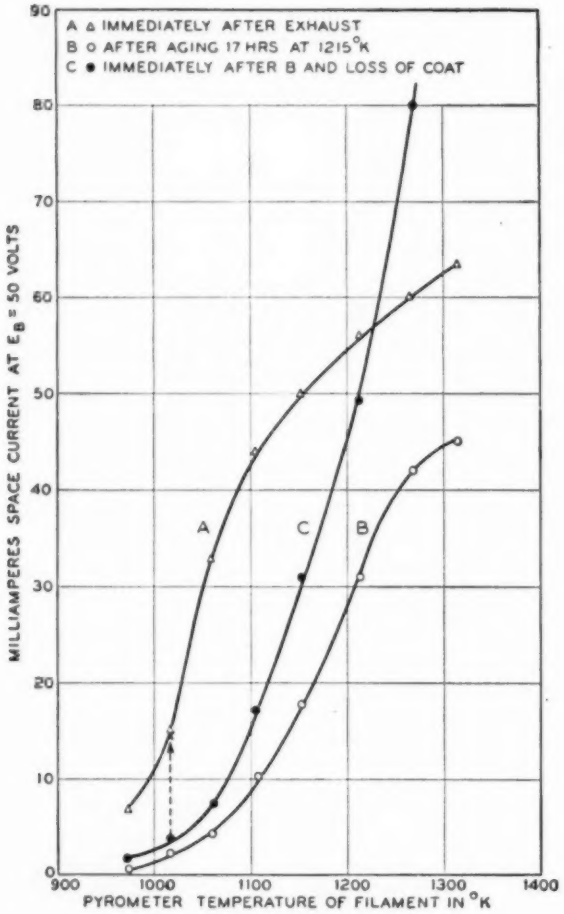


FIG. 10. Tube PE185. Effect of removing coating on thermionic emission. Emission is *increased* immediately following loss of coating from Konal filament.

whom the manuscript was discussed. Support and advice from J. R. Wilson were invaluable.

APPENDIX A

Loss of Coat Does not Usually Decrease Thermionic Emission Immediately

Figure 10 shows that loss of coating from the Konal filament did not decrease emission; thermionic quality was tested immediately before and immediately after the coating dropped off. This was possible because the immediate occasion of poor adherence is invariably a large temperature change such as cooling from operating to ambient. Figure 10 shows a small *increase* in activity (about fivefold).

Becker and Sears[18] attributed a 10^4 fold emission decrease to loss of coating, but their filament was aged 17 hours at a high temperature between their first emission test and loss of coat which was then followed by their second emission test. Jones[19] reported briefly a result which partially confirmed Becker and Sears. Such results are extremely unusual. The writers have observed loss of coating from hundreds of cathodes on life test, but have never observed an immediate decrease in emission. When the coating adheres, gradual decline of emission is usual; decline is greatly accelerated by loss of coat and commerical failure usually occurs a few hundred hours later. Figure 10, typical of results when coating is first lost, is included to indicate the behavior observed in substantially all cases.

[18] J. A. Becker and R. W. Sears, Phys. Rev. **38**, 2193 (1931).
[19] T. J. Jones, *Thermionic Emission* (Methuen and Company, Ltd., London, England 1936), p. 82.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Forward Characteristic of Germanium Point Contact Rectifiers

MELVIN CUTLER
Hughes Aircraft Company, Semiconductor Laboratory, Culver City, California
(Received November 15, 1954)

A quantitative discussion of the forward characteristic is given for point contact rectifiers. It shows that the theoretical value of $d\ \ln I/dV$ in the exponential region is nearly q/kT, and that the experimental slopes which are a fraction of this value cannot be the result of the spreading voltage. Recombination of injected carriers has little effect except at large currents beyond the exponential region. Relations are derived for the limit of very large currents, which show that γ approaches a value smaller than $(1+b)^{-1}$, the hole concentration at the contact approaches a maximum value proportional to the equilibrium concentration of holes at the surface, and the spreading voltage becomes a linear function of the current.

I. INTRODUCTION

THE theory for the behavior of point contact rectifiers has, for a long time, been in poor agreement with experiment.[1-4] This paper is concerned with the theory of the forward characteristic, with reference to *n*-type germanium rectifiers. For the purpose of discussion, the forward-current curve is divided into several parts: the small current region, where a simple theory is applicable[5,6]; the exponential region, where the plot of $\log I$ as a function of V is approximately linear; and a so-called "ohmic" region, at large voltages, where the current changes relatively slowly with voltage, although it may not actually be linear. There will also be occasion to refer to a transition region between the exponential region and the ohmic region.

The ohmic region was originally ascribed to a linear spreading resistance such as could be expected with a small area contact on a metal.[1] The anomalously low value observed for the differential resistance was later explained qualitatively as the effect of injection of holes, with a corresponding decrease in the resistivity.[3]

The exponential region is still imperfectly understood. Emission theory for small currents through the surface barrier requires that $d\log I/dV$ have a value of q/kT. Experimental curves have slopes of the order of one-half of this value. The value varies with different diodes, and in some cases the $\log I-V$ curve seems to have more than one straight line segment.[1,4]

The recent recognition of the nonohmic nature of the spreading resistance in the range of currents usually measured has led to an improvement in the understanding of diode behavior. Banbury has corrected the $I-V$ characteristic by adding a nonlinear spreading voltage calculated on the assumption that the hole current fraction γ is equal to unity.[7] Recombination was assumed to be of negligible importance.

More recently, Swanson has given a semiquantitative treatment of diode theory which is improved by removal of the assumption that γ is unity.[8] The barrier equation for electrons is corrected for the effects of hole injection, but the hole current is assumed to be limited by diffusion, as in small-current theory. Again, recombination of injected holes is assumed to have negligible effect. The results indicate that γ approaches $(1+b)^{-1}$ at large currents, that the hole concentration increases without limit, and that the current approaches, approximately, a quadratic function of the voltage. Swanson notes that the equation for hole current through the barrier cannot be accurate when the potential approaches the equilibrium barrier height.

We shall present a more quantitative discussion of the theory. The approach which is used is to combine formally the flow equations for the barrier region, derived from emission theory,[6] with the corresponding equations for the flow of carriers in the region beyond the barrier. The latter are referred to as the spreading or diffusion equations. Numerical solutions of the combined equations are obtained for typical values of the contact parameters, and limiting relationships are derived for very large currents.

The solution is more accurate than that of Swanson, since diffusion is not assumed necessarily to control the flow of holes in the barrier region. The removal of this limitation is particularly important when the contact radius is small, or when the current is large. It leads to conclusions for the limit of very large currents which differ from those derived by Swanson: γ approaches, in all cases, a value which is smaller than rather than equal to $(1+b)^{-1}$. The hole concentration approaches a value which depends on the height of the barrier at equilibrium, and the current becomes a linear function of the voltage, rather than quadratic.[9]

An important question to be considered is whether

[1] R. C. Torrey and C. A. Whitmer, *Crystal Rectifiers* (McGraw-Hill Book Company, Inc., New York), Chap. 4.
[2] S. Benzer, J. Appl. Phys. **20**, 804 (1949).
[3] J. Bardeen and W. H. Brattain, Phys. Rev. **75**, 1208 (1949).
[4] H. J. Yearian, J. Appl. Phys. **21**, 214 (1950).
[5] J. Bardeen, Bell System Tech. J. **29**, 469 (1950).
[6] M. Cutler, Phys. Rev. **96**, 255 (1954).
[7] P. C. Banbury, Proc. Phys. Soc. (London) **66B**, 833 (1953).
[8] J. A. Swanson, J. Appl. Phys. **25**, 314 (1954).
[9] Although Swanson's main discussion indicates that his conclusions are as stated above, he does predict a situation at "very high currents" which is in qualitative agreement with our results. He seems to feel that this is associated with a breakdown in the validity of the Boltzmann distribution, whereas our results show that only the diffusion assumption is at fault. The correct equations lead to a description which is complete within the limitations of the other assumptions (no heating, and electrical neutrality).

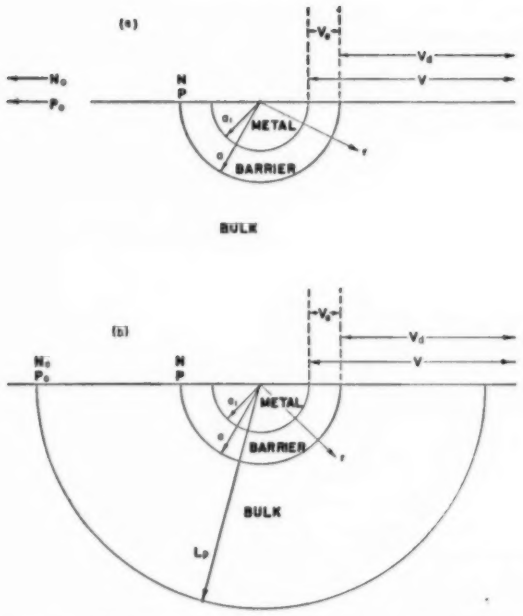


FIG. 1(a). Model for point contact theory, assuming no recombination where γ is constant for all values of radial distance r. (b) Model for point contact theory, assuming recombination occurs entirely at $r=Lp$ where γ is constant for $a<r<Lp$, and is zero for $r>Lp$.

the observed slope of log$I$ *vs* V can be accounted for by the effects of the spreading voltage. This was indicated by Banbury, and the question was left open by Swanson. The solutions derived here, assuming no recombination, give a negative answer. In order to remove the possibility that recombination may affect this conclusion, solutions of the diffusion equations are obtained for a case where the hole lifetime is finite, and smaller than the value for most experimental situations.

II. SOLUTION OF EQUATIONS

The model used for the point contact is shown in Fig. 1(a). The metal contact is hemispherical, and has a radius a'. The surface barrier has a thickness $(a-a')$, and the boundary between the barrier region and the bulk of the semiconductor is assumed to be at the radial distance $r=a$. In most cases of interest, the barrier thickness is small compared to the radius of the metal contact, so that the latter is essentially the same as a. We shall not, therefore, consider the effects on the value of a of variation in the barrier thickness with applied potential.

The barrier has an equilibrium height ϕ, and is assumed to be homogeneous and to behave ideally; that is, an applied potential V_e between the metal and the point $r=a$ has the effect of reducing the height of the barrier by this amount. In all of the discussions which follow, electrical neutrality is assumed: the difference between the concentrations of electrons and holes $(n-p)$ at $r=a$ is constant and is equal to the equilibrium value n_0-p_0. Also, $\nabla(n-p)=0$. The solution for the diffusion equations which is used in this section contains the assumption that the hole lifetime τ_p is infinite. The effect of recombination is considered later.

The equations for the flow of holes and electrons through the barrier region are[6]:

$$I_p/I_{pe}=\exp(qV_e/kT)-p/p_0, \tag{1}$$

$$I_n/I_{ne}=(n/n_0)\exp(qV_e/kT)-1, \tag{2}$$

where $I_{pe}=qv2\pi a^2p_0$, $I_{ne}=qv2\pi a^2n_0\exp(-\phi/kT)$, and v is the thermal velocity of the carriers. After using the electrical neutrality assumption to eliminate n, one can eliminate V_e between the two equations to obtain an equation for the barrier region which relates I, γ, and p:

$$I=(1+p^*)(p^*/p_0^*)\times[(1-\gamma)I_{ne}^{-1}-\gamma(1+p^*)I_{pe}^{-1}]^{-1}. \tag{3}$$

We have used the reduced variable $p^*=(p-p_0)(n-n_0)$, and p_0^* denotes $p_0/(n_0-p_0)$. It has been assumed that $p_0^*\ll 1$. This limits the discussion to extrinsic semiconductors, although it implies no restriction on the method of analysis. It is to be noted that Eq. (3) differs from the one used by Swanson in that the latter does not contain the term $\gamma(1+p^*)I_{pe}^{-1}$.

The solution of the diffusion equations has been derived previously.[8] For purposes of later discussion, we present here the differential equations from which they are derived, as well as the equations of the final solution:

$$[b\gamma+p^*(b\gamma-1+\gamma)]^{-1}\times(1+2p^*)dp^*/dr=-Ia/I_{nd}r^2, \tag{4a}$$

$$-(b\gamma-1+\gamma)Ia/I_{nd}r^2=-(q/kT)dV_d/dr+2dp^*/dr, \tag{5a}$$

where $I_{nd}=qD_n2\pi an_0$, D_n is the diffusion constant for electrons, and V_d is the voltage drop in the diffusion region, or the spreading voltage. [See Fig. 1(a).] In contrast to the main body of the discussion, V_d and p^* in the aforementioned equations are functions of r rather than their values at $r=a$. This should cause no confusion, since V_d and p^* are regarded as functions of r only in equations in which r appears. Here γ is constant since we assume that no recombination occurs.

When Eqs. (4a) and (5a) are integrated between $r=a$ and $r=\infty$, one obtains the equations which describe the situation in the diffusion region (where now V_d and p^* refer to the values at $r=a$):

$$(b\gamma-1+\gamma)I/I_{nd}=2p^*+[(b\gamma+1-\gamma)/(b\gamma-1+\gamma)]\times\log[1+p^*(b\gamma-1+\gamma)/b\gamma], \tag{4}$$

$$V_dq/kT=2p^*-(b\gamma-1+\gamma)I/I_{nd}. \tag{5}$$

At $r=a$, the boundary between the barrier region and the bulk region, (or diffusion region), one has common values for p, n, I_p, and I_n (or the corresponding variables p^*, I, and γ). The total voltage V is the sum

of V_e (the drop in the barrier) and the spreading voltage V_d. One can therefore obtain the complete solution by combining Eqs. (3), (4), (5), and (1) or (2). This cannot be done analytically, so that numerical solutions are constructed by the following procedure. A set of values are chosen for a, ϕ, and n_0 to determine values for all the parameters in the equations, that is, I_{nd}, I_{pe}, I_{ne}. Equations (3) and (4) are combined to find the dependence of γ and p^* on I. Then, for each value of I, the corresponding values of V_e and V_d are obtained from Eq. (1) or (2) and Eq. (5), respectively. Thus, one obtains the curve for the total voltage $V(I)$, for any particular set of parameters, as well as $\gamma(I)$ and $p^*(I)$.

Numerical solutions were obtained in this way for several sets of point contact parameters. The parameters are believed to cover a range in which typical diodes are likely to fall. They lie in the following ranges: a, from 1.5×10^{-4} to 1.6×10^{-3} cm; ϕ, from 0.09 ev to 0.40 ev; (n_0-p_0), from 3×10^{14} cm^{-3} to 7×10^{14} cm^{-3}. (It can be shown that the limited range in n_0 does not imply a restriction on the generality of the results.)

Some typical sets of curves are shown in Figs. 2 and 3. Although I is the natural independent variable, the curves are plotted as functions of V, in keeping with convention. It is noted that the value of $d\log I/dV$ in the exponential region is approximately q/kT. We shall discuss the curves in more detail later.

III. LIMITING RELATIONS

It is unfortunate that the combined solution can only be constructed by numerical methods, so that one cannot describe its properties by means of analytical expressions. However, it is possible to deduce some interesting relationships in the limit where I/I_{nd} becomes very large.

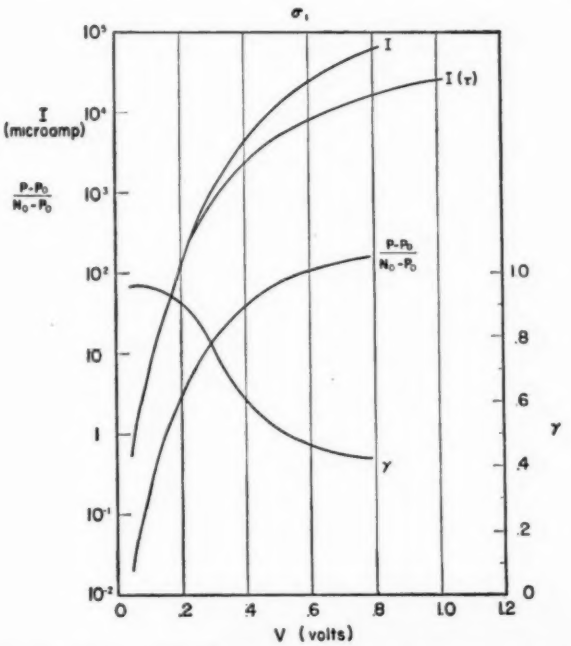


FIG. 2. Computed curves for $a=1.64\times10^{-3}$ cm, $\rho=5.0$ ohm cm, $\phi=0.336$ ev; $(p_f/n_0)=5000$. $I(\tau)$ has correction for $\tau_p=7.8$ μsec.

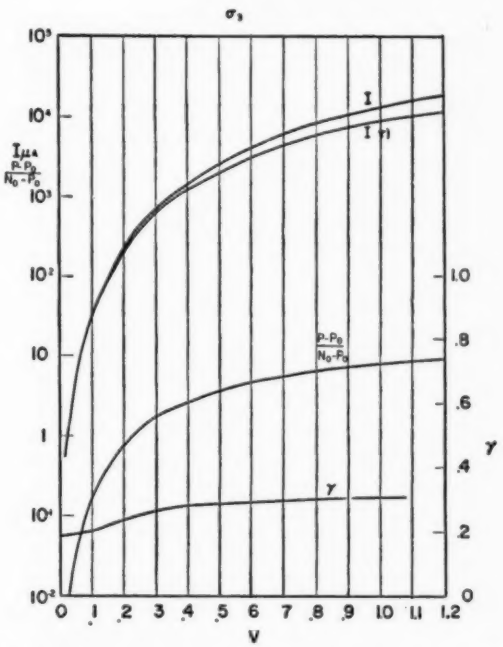


FIG. 3. Computed curves for $a=1.64\times10^{-3}$ cm, $\rho=5.0$ ohm cm, $\phi=0.218$ ev; $(p_f/n_0)=50$. $I(\tau)$ has correction for $\tau_p=7.8$ μsec.

On examining Eqs. (3) and (4), it is found that I can consistently approach indefinitely large values in both equations only if p and γ approach limiting values p_f and γ_f, respectively, which satisfy the following relationships:

$$1+[(p_f-p_0)(b\gamma_f-1+\gamma)/b\gamma_f(n_0-p_0)]=0. \quad (6)$$

$$(1-\gamma_f)I_{ne}^{-1}-\gamma_f I_{pe}^{-1}[1+(p_f-p_0)/(n_0-p_0)]=0. \quad (7)$$

The solutions for p_f and γ_f are:

$$(p_f-p_0)/(n_0-p_0)=bI_{pe}/I_{ne}, \quad (8)$$

$$\gamma_f=[1+b+I_{ne}/I_{pe}]^{-1}. \quad (9)$$

Since $I_{pe}/I_{ne}=(p_0/n_0)\exp(\phi/kT)$, it is apparent that the hole concentration near the contact approaches a maximum value, which is equal to $bp_0\exp(\phi/kT)$. Aside from the factor b, this might be expected in a rough way on the basis of p-n junction theory, since $p_0\exp(\phi/kT)$ is the concentration of holes at the top of the surface barrier or the surface "p-region." (We shall also refer to the surface hole concentration as p_s.)

Unless the surface barrier is small, I_{ne}/I_{pe} will be small compared to unity, so that the limiting value of γ will be somewhat less than $(1+b)^{-1}$. Otherwise, γ_f will be considerably smaller than $1/(1+b)$. The behavior of γ is quite similar to Swanson's result. But the fact that γ_f is less than $(1+b)^{-1}$ rather than equal to it leads to the significant difference that the current

is a linear function of the voltage at very large currents, rather than quadratic.

The limiting spreading resistance R_s is obtained by substituting Eqs. (8) and (9) in Eq. (5). As I/I_{nd} becomes large, the term $2p^*$ in Eq. (5) approaches a constant value, so that

$$(dV_d/dI)_{I=\infty} = -(q/kTI_{nd})(b\gamma_f - 1 + \gamma_f), \quad (10)$$

or, for a resistivity ρ,

$$R_s = (\rho/2\pi a)(I_{ne}/I_{pe})[b+1+I_{ne}/I_{pe}]^{-1}. \quad (11)$$

Since, in the limit of large I, dV_s/dI is roughly proportional to I^{-1} and therefore approaches zero, no contribution to the impedance can be expected from the barrier region.

Equation (11) indicates that if $I_{ne} \ll I_{pe}$, the limiting spreading resistance is smaller, as the result of injection, by the factor $I_{ne}/I_{pe}(b+1)$. This, obviously, is simply a correction in the conductivity for an ohmic spreading resistance where both n and p are equal to p_f. We must caution here that Eq. (11) may be inaccurate if recombination occurs. This factor will be discussed in detail below.*

IV. THE EFFECT OF RECOMBINATION

Hole recombination, such as would occur in practical point contacts, will not affect the theoretical behavior of $p(I)$ and $\gamma(I)$. This is proved rigorously by considering the accuracy of Eq. (4) in the case of finite lifetime τ_p. Equation (4) is the solution for $\tau_p = \infty$, and was obtained by integrating Eq. (4a) with respect to the radial distance r from $r=a$ to $r=\infty$, with the boundary condition that $p(r=\infty)=p_0$ (or $p^*(r=\infty)=0$). If one considers an arbitrary model illustrated by Fig. 1(b), where all the injected holes recombine instantly on reaching the distance $L_p=(D_p\tau_p)^{\frac{1}{2}}$, then the boundary condition is $p^*=0$ at $r=L_p$. The solution is Eq. (4) modified only in that it contains an additional factor $(1-a/L_p)$ on the left-hand side. The exact solution must lie between these two, for the holes are swept farther than L_p, on the average, before recombining. Therefore, Eq. (4) is accurate if $a \ll L_p$. Since this is the usual experimental situation, the accuracy of the solution for γ and p^* as a function of I, derived from Eqs. (1), (2), (3), and (4), will not be affected by neglect of recombination.

The same type of reasoning, applied to the derivation of the spreading voltage in Eq. (5) must contain an added consideration: In the second model [Fig. 1(b)], one must add a term for the voltage drop for the flow of current from $r=L_p$ to $r=\infty$, through material of normal resistivity. The term $(I/I_{nd})(a/L_p)$ should then be added to the right side of Eq. (5), (as well as the negligible factor $(1-a/L_p)$ in the second term). Therefore, according to the second model, when $(p_f-p_0)/(n_0-p_0) \gg 1$,

$$R_s = (\rho/2\pi a)\{[I_{ne}/(b+1)I_{pe}]+a/L_p\}. \quad (12)$$

Recombination will have an appreciable effect, then, if a/L_p is not small compared to I_{ne}/I_{pe}; this may well occur in practice. Equations (11) and (12) bound the actual differential resistance at large currents, since the former neglects recombination, and the latter exaggerates its effect.

This leaves open the question of whether the spreading potential can account for the observed values of $d\log I/dV$ in the exponential region of the diode characteristic. The effect of recombination would be to increase V for any value of I. It is conceivable, then, that the correction will give better agreement with experiment in the exponential region.

In order to check this possibility, numerical solutions of the diffusion equations with recombination were obtained with the use of analog computing machines. They were obtained for the particular case where $L_p=10a$, which corresponds to rather small values of τ_p (0.05 to 8 microseconds). Equations (4a) and (5a) were used in the machine solution, where γ was now regarded as a function of r. The additional equation, which serves to determine the behavior of γ, is obtained from the conservation law:

$$(d\gamma/dr)(I/2\pi qr^2) = (p_0-p)/\tau_p. \quad (13)$$

The computed $I-V$ curves, corrected for recombination effects, are shown in Figs. 1 and 2, marked $I(\tau)$. It is found that the correction is noticeable only in the very high current region approaching the ohmic region, beyond the exponential region. It is therefore not expected that the spreading voltage can account for observed slopes in the exponential region of diode characteristics, whether or not recombination occurs.

V. DISCUSSION

Some computed curves for $\log I$, $\log[(p-p_0)/(n_0-p_0)]$, and γ are plotted as a function of V in Figs. 1 and 2. The behavior of $\log[(p-p_0)/(n_0-p_0)]$ and γ is consistent with the predicted limiting relations for very small and very large values of I/I_{nd}. The exponential region is characterized by $V_d \ll V_e$. Here, γ is nearly constant, and $(p-p_0)$ increases rapidly with increasing current. In the intermediate region which follows, $V_d \sim V_e$, γ changes rapidly, and $(p-p_0)$ increases less rapidly. Following this is the ohmic range, in which $V_d \gg V_e$. In this range, $p-p_0$ and γ change relatively slowly. The reason is largely that they depend on I, and I is increasing more slowly with V.

The dependence of I on V in the "ohmic" range is complicated by the fact that γ is changing as well as I. As a result, the $V-I$ relationship is difficult to analyze. The computed V_d-I curves seem to show a power

Note added in proof.—It has been pointed out to the author by K. Lehovec that Eq. (1) cannot be valid when the barrier disappears, i.e., where $qV_s > \phi$. Therefore, the limiting relations must be regarded as having only mathematical significance. Clearly, then, Eq. (8) cannot be physically valid, since p must remain smaller than p_s. Whether the value of γ will become smaller than $(1+b)^{-1}$ in the range of validity, and R_s becomes constant, will depend on the relative values of a, ϕ, and ρ.

relationship $I\alpha V_d^{\mu}$, where μ has values ranging between 1.0 and 3.6. This is qualitatively different from Swanson's result, and it is consistent with experimental data.[8] Swanson's results predict that $\gamma\rightarrow(1+b)^{-1}$, so that μ is exactly two, and that deviations occur because the electrical neutrality assumption is not accurate. Although the latter may be true, the value of μ is not necessarily 2.0 because γ is varying. At larger currents, γ will pass below $(1+b)^{-1}$, and μ will revert to the value unity, although it may occur beyond the usual experimental range.

The behavior of $V_d(I)$ also leads to the result that $d\ln I/dV\cong q/kT$. Banbury showed that if γ is unity, one will obtain a region of the current characteristic where both $\exp(qV_e/kT)$ and $\exp(qV_d/kT)$ are proportional to I, so that $d\ln I/dV=q/2kT$. Similarly, Swanson's results show that if $\gamma>(1+b)^{-1}$ and is constant over a large range of I, $d\ln I/dV<q/kT$. However, when the equations are treated quantitatively, it is found that γ starts to decrease when $I>I_{nd}$, while $V_d\ll V_e$. When the point is reached where $V_d\sim V_e$, the behavior of γ is such that V_d is not proportional to $\ln I$, but to some power of I. As a consequence, the current characteristic changes abruptly from the "classical" exponential curve with $d\ln I/dV\cong q/kT$ to one in which $I\alpha V^{\mu}$. As previously noted, recombination does not alter the situation qualitatively.

One is therefore led to the conclusion that the impedance beyond the barrier does not account for the behavior of diodes in the exponential region of the forward characteristic. The reasons must be sought in the model for the flow of carriers through the barrier. This discrepancy between theory and experiment may be an extension of the situation in the small current region: The value of $I[\exp(qV/kT)-1]^{-1}$ for small V (say, $|V|<0.1$ volt) is not constant, but generally decreases with increasing V.[6] This function, which is theoretically $(I_{ne}+I_{pd})$, is generally large compared to the computed value of I_{pd}; it apparently indicates a low value for γ at small currents.

One possible way to look at the situation is to regard I_{ne} as a variable. This is consistent with some proposed models for barrier flow; in the multicontact hypothesis the apparent barrier height ϕ is increased as the bias is increased, because the low barrier paths are made less effective by the associated IR drop.[10] Thus, I_{ne} apparently decreases with increasing forward bias. A slope $d\ln I/dV=(1-\alpha)q/kT$ corresponds to $\phi=\text{const}+\alpha V_e$. This would cause an increase in γ on top of the changes which would occur for constant I_{ne}.

Recent work indicates that the explanation may lie in surface leakage.[11,12] The effect is analogous to the multicontact theory, except that the effective area as well as the barrier may be varying with bias. Contrary to what one might suppose at first glance, the leakage current can affect the forward as well as the reverse characteristic, as long as the impedance of the surface path is not large compared to that of the barrier under the metal contact.

One could obtain more definite information about the behavior of the barrier region if, in addition to $\gamma(I)$, the simultaneous behavior of $\gamma(I)$ or $p(I)$ could be measured with sufficient accuracy. In such a case, it is possible to work back through the use of Eqs. (4) and (5) to obtain the empirical behavior of I_p and I_n as a function of V_e and p.

In previous work, the approximation

$$p/p_0=\exp(qV_e/kT)$$

has often been used, and it seems worthwhile to discuss its limitations. It has been shown for small current theory[6] that the accuracy of the diffusion control approximation depends on the relation $I_{pe}/I_{pd}\gg 1$, or $va/D_p\gg 1$. This ratio has approximately the value 60 for $a=10^{-3}$ cm. As I increases above I_{nd}, the value of I_p/I_{pe} in Eq. (1) becomes larger compared to p/p_0. When it becomes comparable, the diffusion control approximation is inaccurate. The point at which this occurs depends on the magnitude of va/D_p (which is approximately the ratio of p/p_0 to I_p/I_{pe} at small currents) and on the value of p_f. When p approaches the limit p_f, it is obvious that I_p/I_{pe} must gain on p/p_0. For $a\sim 10^{-3}$ cm, the approximation seems to be good up to the range of high currents (in the ohmic range). But if a is smaller, appreciable errors occur even in the transition range, where V_d first becomes noticeable. Of course, as we have noted, it is necessary to use the accurate hole emission equation in order to obtain the correct relations in the limit of very large currents.

A linear relationship between I and V at very large currents has been predicted on the basis of an appreciable deviation from the Boltzmann distribution at a large forward bias. A careful consideration of the situation indicates that there is no necessary correlation between the disappearance of the barrier and a deviation from the Boltzmann distribution. A trivial example is the case where the barrier is very small; there is no rectification, and Ohms' law applies. Confusion may arise from the fact that the emission equations are derived on the assumption that the carriers cross the barrier at thermal velocity. However, they come practically to equilibrium (or quasiequilibrium) with other carriers within a few free-flight distances from the barrier; the situation is not altered by the disappearance of the barrier. Indeed, the assumption that carriers move with thermal velocity over a distance of the order of the scattering distance can also be used to derive the Smoluchowski equations, upon which the spreading equations are based.[6]

In summary, a quantitative treatment for the theory of diode behavior has been derived. The disagreement between the theoretical and the experimental current characteristic in the exponential region cannot be removed by a detailed consideration of the spreading

[10] Johnson, Smith, and Yearian, J. Appl. Phys. **21**, 283 (1950).
[11] M. Cutler and H. M. Bath, J. Appl. Phys. **25**, 1440 (1954).
[12] J. B. Gunn, Proc. Phys. Soc. (London) **67B**, 409 (1954).

equations. It must be assumed, then, that the model for the barrier is not accurate.

In spite of this limited agreement, some conclusions, which are of general value, can be drawn for behavior at large currents: In the limit of very large currents, the voltage must approach a linear function of the current, although the usual experimental ohmic range may be an intermediate one where a power law applies. Recombination can have an appreciable effect on the spreading resistance, for which limiting relations have been derived. The hole concentration at the contact approaches a maximum value equal to $bp_0 \exp(\phi/kT)$, and at the same time γ approaches a limit which is smaller than $(1+b)^{-1}$.

The author wishes to express his appreciation for the assistance he received through discussions with members of this Laboratory, and through the calculations carried out by computer groups in this Company. He wishes to thank particularly I. Pfeffer for making difficult machine computations.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Similarity Solution for a Spherical Shock Wave

RICHARD LATTER
The Rand Corporation, Santa Monica, California
(Received February 28, 1955)

The point-source, spherical shock wave moving into a constant density γ-law gas is considered in the limit of infinite shock strength from the point of view of the Richtmyer-von Neumann viscosity technique. A similarity solution of this problem is shown to exist and is obtained for various boundary conditions with $\gamma=1.4$. The solutions are obtained analytically in that part of the flow field not involving viscosity, and numerically in the other parts of the flow field. It is found that whereas all discontinuities of the physical parameters are removed by the viscosity there remain discontinuities in the slopes of these parameters at the shock front. It is indicated, moreover, that the complete flow field depends upon the form and magnitude of the viscosity.

I. INTRODUCTION

THE existence of shock waves in a hydrodynamic flow field introduces free-boundary discontinuities into the physical parameters of the system. Such discontinuities cause considerable analytic as well as numerical complications for the treatment of hydrodynamic problems. A means for avoiding these difficulties, particularly for numerical calculations, has been developed by Richtmyer and von Neumann.[1] They observed that the addition of a particular viscosity-like term into the hydrodynamic equations could lead to continuous shock-flow fields in which the discontinuities at the shock wave were removed and replaced by a region in which the physical parameters changed rapidly, but smoothly. Moreover, the change of the physical parameters through this transition region was, for plane flow fields, in agreement with the Rankine-Hugoniot relations.

Richtmyer and von Neumann obtained an analytic solution of a flow field including the artificial viscosity only for the case of a plane shock wave. In the present discussion, it is shown that in the limit of infinite shock strength, a point-source spherical shock flow which includes the artificial viscosity has a similarity solution for a γ-law gas. The latter has been obtained by numerical and analytic integration of the ordinary differential equations resulting from the similarity of the flow.

An important consequence of the similarity solution described here is that it reveals that though the discontinuities of the physical parameters are removed from the flow field by the viscosity, the rates of change of these parameters have discontinuities. Moreover, the flow field itself is modified by the viscosity, approaching the Taylor[2] strong-shock field only in the limit of vanishing viscosity.

II. FORMULATION

When the viscosity term suggested by Richtmyer and von Neumann is included into the hydrodynamic equations for a spherically symmetric flow in a γ-law medium, the hydrodynamic equations in Eulerian form become

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial r}=-\frac{1}{\rho}\frac{\partial(p+q)}{\partial r}, \tag{1}$$

$$\frac{\partial \rho}{\partial t}+u\frac{\partial \rho}{\partial r}=-\rho\left(\frac{\partial u}{\partial r}+\frac{2u}{r}\right), \tag{2}$$

$$\frac{\partial p}{\partial t}+u\frac{\partial p}{\partial r}=\frac{\gamma p+(\gamma-1)q}{\rho}\left(\frac{\partial \rho}{\partial t}+u\frac{\partial \rho}{\partial r}\right), \tag{3}$$

where p is the static pressure, ρ the density of the medium, u the material velocity, q the artificial viscosity, and (r,t) the space-time Eulerian coordinates.

[1] R. D. Richtmyer and J. von Neumann, J. Appl. Phys. **21**, 232 (1950).

[2] G. I. Taylor, Proc. Roy. Soc. (London) **A201**, 159 (1950).

The caloric equation of state of the medium is assumed to be

$$\epsilon = p/(\gamma-1)\rho, \tag{4}$$

where ϵ is the internal energy per unit mass and $\gamma>1$. The adiabatic equation of state is assumed to be

$$p/\rho^\gamma = \sigma(S), \tag{5}$$

where $\sigma(S)$ is a function of the entropy S only. The quantity q in Eqs. (1) to (3) is the viscosity term. That dependence of q on p,ρ, and u which prohibits discontinuities in the physical parameters is not uniquely prescribed by Richtmyer and von Neumann. The form for q adopted for the present discussion is consistent with their requirements and is

$$q = -\frac{1}{2}K^2\rho r^2\frac{\partial u}{\partial r}\left(\left|\frac{\partial u}{\partial r}\right| - \frac{\partial u}{\partial r}\right). \tag{6}$$

The arguments presented here are not restricted to the form of q in Eq. (6), but can be generalized to include other possibilities. However, some of the detailed results depend upon the specific form of q which is used; in particular, not all forms for q admit of similarity solutions.

The possibility of a similarity solution of Eqs. (1) to (3) and (6) is suggested by the calculations of Taylor[2] on the point-source, strong-shock problem, in which the shock is treated as a discontinuity. It may be observed from Eq. (6) that the Taylor equations and the present ones are identical in the region between the origin and the Taylor shock front since $q=0$ in this region, as follows from the property $\partial u/\partial r>0$ for the material velocity behind the Taylor strong shock. This observation indicates that the Taylor similarity variables may be suitable also for the present problem. Thus the Eqs. (1) to (3) and (6) are examined for a similarity solution of the form

$$\frac{p(r,t)}{\rho_0} = \frac{f(x)}{R^\beta}, \tag{7}$$

$$\frac{\rho(r,t)}{\rho_0} = \psi(x), \tag{8}$$

and

$$u(r,t) = \phi(x)/R^\alpha, \tag{9}$$

where $x=r/R$, R is a function of time only, which at present need have no relation to the shock position, ρ_0 is an arbitrarily prescribed constant with the dimensions of density, and α and β are constants yet to be determined. To simplify the subsequent calculations the result $\beta=2\alpha$, which may be easily derived from the arguments to follow, is assumed without proof. The expressions in Eqs. (7) to (9) are now substituted into Eqs. (1) to (3) and (6), using the relationships, which result from the change of independent variables from (r,t) to (x,t),

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t} - \frac{x\dot{R}}{R}\frac{\partial}{\partial x},$$

$$\frac{\partial}{\partial r} = \frac{1}{R}\frac{\partial}{\partial x},$$

where $\dot{R}=dR/dt$. Equation (1) then becomes in terms of the new variables

$$-\alpha\phi R^\alpha\dot{R} - x\phi' R^\alpha\dot{R} + \phi\phi' + \frac{f'+g'}{\psi} = 0, \tag{10}$$

where the prime indicates differentiation with respect to x. The quantity g is related to q by the equation

$$\frac{q}{\rho_0} = \frac{K^2}{2R^{2\alpha}}\psi x^2\phi'(|\phi'|-\phi') = \frac{g(x)}{R^{2\alpha}}, \tag{11}$$

which arises directly from Eq. (6) using Eqs. (7) to (9). If, now, it is assumed that

$$R^\alpha\dot{R} = A, \text{ a constant}, \tag{12}$$

then Eq. (10) becomes an ordinary differential equation in the independent variable x, namely,

$$-A(\alpha\phi + x\phi') + \phi\phi' + \frac{f'+g'}{\psi} = 0. \tag{13}$$

Quite similarly, Eqs. (2) and (3) become, respectively,

$$-Ax\psi' + \phi\psi' + \psi\phi' + 2\frac{\phi\psi}{x} = 0 \tag{14}$$

and

$$-A2\alpha f + (\phi - Ax)f' - \frac{\gamma f + (\gamma-1)g}{\psi}(\phi - Ax)\psi' = 0. \tag{15}$$

The constant α is determined by requiring a constant energy in the hydrodynamic flow. If E is the total energy, then

$$E = 4\pi\int_0^\infty r^2 dr(\rho u^2/2 + p/(\gamma-1)),$$

which becomes in terms of the present similarity variables

$$E = 4\pi\frac{R^3}{R^{2\alpha}}\rho_0\int_0^\infty x^2 dx(\psi\phi^2/2 + f/(\gamma-1)). \tag{16}$$

Since R is a function of time, this equation requires that $\alpha=\frac{3}{2}$.

If new dimensionless quantities

$$F = \frac{f}{A^2}, \quad G = \frac{g}{A^2}, \quad \text{and} \quad \Phi = \frac{\phi}{A}$$

are introduced into Eqs. (11) and (13) to (16), these

equations read

$$G=\frac{K^2}{2}x^2\psi\Phi'(|\Phi'|-\Phi'), \tag{11'}$$

$$-\left(\frac{3}{2}\Phi+x\Phi'\right)+\Phi\Phi'+\frac{F'+G'}{\psi}=0, \tag{13'}$$

$$(\Phi-x)\psi'+\psi\Phi'+2\psi\frac{\Phi}{x}=0, \tag{14'}$$

$$-3F+(\Phi-x)F'-\frac{\gamma F+(\gamma-1)G}{\psi}(\Phi-x)\psi'=0, \tag{15'}$$

and

$$E=4\pi\rho_0A^2\int_0^\infty x^2dx(\psi\Phi^2/2+F/\gamma-1). \tag{16'}$$

The existence of a solution of Eqs. (13′) to (15′) for a diverging flow from a point source with a shock discontinuity at $x=1$ has been demonstrated by Taylor. But for such a diverging flow $d\Phi/dx>0$, so that $G=0$, and therefore the Taylor solution constitutes also a solution of the present problem in the region from $x=0$ to $x=1$, where R is identified with the position of the Taylor shock front. In the Taylor solution the quantities F, Φ, and ψ at $x=1$ are consistent with the Rankine-Hugoniot conditions for a strong shock. The presence of viscosity removes the discontinuities in the physical quantities, and consequently, the Rankine-Hugoniot conditions have no special significance in the viscosity formalism. However, if the viscosity is to leave the flow field unaltered, then the Rankine-Hugoniot conditions should apply across the shock transition region wherein the viscosity is effective. In a similarity problem of the present type it is clear that this situation is not possible since the width of the transition region will be a constant with respect to the similarity variable $x=r/R$. The transition region will therefore contain an amount of mass given by

$$m(r,t)=4\pi\int r^2dr\rho(r,t),$$

which in terms of the similarity variables is

$$m(r,t)=4\pi\rho_0R^3\int x^2dx\psi(x).$$

The limits of integration are fixed by the width of the transition region and depend, consequently, only upon x. The time dependence of R given in Eq. (12) implies then that the mass, $m(r,t)$, within the transition region varies with time. For the validity of the Rankine-Hugoniot conditions across the transition region, $m(r,t)$ would have to be a constant, but this condition has been proved to be violated. This impossibility of meeting the Rankine-Hugoniot conditions, is, of course, a property of the assumed type of similarity expressed by Eqs. (7) to (9) as well as of the specific form for q of Eq. (6). The restrictions on the form of q leading to the Rankine-Hugoniot conditions are still not known in general and certainly need further exploration.

Before studying the behavior of the differential equations, it is convenient to formulate the boundary conditions which these equations must satisfy. Namely, all physical quantities must be bounded and continuous. While the artificial viscosity is not a physical quantity, the spirit of the Richtmyer-von Neumann formalism requires treating it as such and therefore q is also assumed continuous. For the presently considered diverging flow field, the physical quantities must satisfy the additional conditions that in the neighborhood of $x=0$, $\Phi=0$, $F\geqslant 0$, $G=0$, and $\psi\geqslant 0$, and in front of the flow the medium is undisturbed so that $\Phi=0$, $F\geqslant 0$, $G=0$, and ψ is a constant. From Eq. (7) it is readily observed that the similarity of the flow requires $F=0$ for large x in the undisturbed medium.

The treatment of Eqs. (11′) and (13′) to (15′) will be divided into the consideration separately of the regions wherein the viscosity is absent and wherein it is present. The region without viscosity will be treated first and will be solved analytically. The region with viscosity will then be explored and joined on by numerical integration to the rest of the flow field.

III. FLOW FIELD WITHOUT VISCOSITY

From the boundary conditions for the diverging flow field, it follows that in a neighborhood of $x=0$, the gradient of the velocity is positive and therefore the viscosity term G is zero. The equations defining the flow then simplify to

$$-\left(\frac{3}{2}\Phi+x\Phi'\right)+\Phi\Phi'+\frac{F'}{\psi}=0, \tag{17}$$

$$(\Phi-x)\psi'+\psi\Phi'+2\psi\frac{\Phi}{x}=0, \tag{18}$$

$$-3F+(\Phi-x)F'-\frac{\gamma F}{\psi}(\Phi-x)\psi'=0. \tag{19}$$

These equations admit of a straightforward analytic integration. Thus by dividing Eq. (18) by $\psi(\Phi-x)$ and by adding the result to Eq. (19), after dividing the latter by $F(\Phi-x)$,

$$\frac{F'}{F}-(\gamma-1)\frac{\psi'}{\psi}+\frac{(x-\Phi)'}{x-\Phi}+\frac{2}{x}=0.$$

This equation integrates directly to

$$x^2(x-\Phi)\frac{F}{\psi^{\gamma-1}}=a. \tag{20}$$

A second integral of Eqs. (17) to (19) may be obtained by multiplying Eq. (17) with $\psi\Phi$ and by adding to Eq. (19) multiplied by $1/(\gamma-1)$. Simplifying the result with the aid of Eq. (18) gives

$$\frac{d}{dx}\left[x^2\left\{F\Phi-(x-\Phi)\left(\frac{F}{\gamma-1}+\frac{1}{2}\psi\Phi^2\right)\right\}\right]=0.$$

Integration of this equation, using the condition $\phi=0$ at $x=0$, yields

$$F\Phi-(x-\Phi)\left(\frac{F}{\gamma-1}+\frac{1}{2}\psi\Phi^2\right)=0. \tag{21}$$

Alternatively, this equation is an immediate consequence of the conservation of energy if this conservation is expressed in integral form. Finally, the third integral is obtained by direct substitution of Eqs. (20) and (21) into Eqs. (17) to (19). Thus, first eliminate F' and ψ' from Eq. (17) using Eqs. (18) and (19); then

$$\Phi'=\frac{\frac{3}{2}\Phi(x-\Phi)(\psi/F)+[3-2\gamma(\Phi/x)]}{\gamma-(x-\Phi)^2(\psi/F)}. \tag{22}$$

Expressing ψ and F in terms of Φ and x gives for Φ'

$$\Phi'=-\frac{\Phi}{x}\frac{2\gamma(\gamma-1)(\Phi/x)^2-3(2\gamma-1)(\Phi/x)+3}{\gamma(\gamma+1)(\Phi/x)^2-2(\gamma+1)(\Phi/x)+2}. \tag{23}$$

This is a standard homogeneous equation and readily integrated giving

$$x=b\frac{\left|\frac{\Phi}{x}-\frac{1}{\gamma}\right|^{(\gamma-1)/(2\gamma+1)}}{\left|\frac{\Phi}{x}\right|^{2/5}\left|\frac{\Phi}{x}-\frac{5}{3\gamma-1}\right|^{(13\gamma^2-7\gamma+12)/5(2\gamma+1)(3\gamma-1)}}. \tag{24}$$

Equations (20), (21), and (24) constitute a complete analytic solution of the differential equations, Eqs. (17) to (19). Since F and ψ are intrinsically positive, the constant a of Eq. (20) is positive.

The behavior of the flow field in the present region is defined by fixing the two arbitrary integration constants a and b. These are in turn specified by requiring definite values for Φ and F at a prescribed position x. It is clear therefore that the Eqs. (17) to (19) describe a two parameter family of flow fields. No further restriction of these parameters is possible before discussing the region containing viscosity.

Only the properties of the present solution relevant to the viscosity formalism will be considered. First, since $\Phi=0$ at $x=0$, Eq. (24) requires that $\Phi/x=1/\gamma$ at $x=0$. Solving Eq. (21) for F/ψ gives

$$\frac{F}{\psi}=\frac{\gamma-1}{2\gamma}x^2\frac{(\Phi/x)^2[1-(\Phi/x)]}{(\Phi/x)-(1/\gamma)}$$

and, moreover, from the positivity of F and ψ, this equation shows that

$$1/\gamma\leqq\Phi/x\leqq 1. \tag{25}$$

Combining Eqs. (24) and (25) leads to the restriction

$$\frac{1}{\gamma}\leqq\frac{\Phi}{x}\leqq\min\left(1,\frac{5}{3\gamma-1}\right). \tag{26}$$

Using this equation with Eq. (23) it is shown immediately that

$$\Phi'>0, \tag{27}$$

for all $x\geqslant 0$ provided that the physical requirement $\gamma>1$ is met. This result is established by observing first that at $x=0$, $\Phi/x=1/\gamma$ and $\Phi'=1/\gamma>0$. If then Φ' is to violate Eq. (27), there must exist a point at which $\Phi'=0$. But this is possible only if

$$2\gamma(\gamma-1)\left(\frac{\Phi}{x}\right)^2-3(2\gamma-1)\frac{\Phi}{x}+3=0.$$

Solving this equation for Φ/x yields

$$\frac{\Phi}{x}=\frac{3(2\gamma-1)\pm[3(4\gamma^2-4\gamma+3)]^{\frac{1}{2}}}{4\gamma(\gamma-1)}.$$

It may be demonstrated analytically from this expression or simply by plotting the two roots as a function of γ that for $\gamma>0$

$$\left(\frac{\Phi}{x}\right)_+>\frac{5}{3\gamma-1}\quad\text{and}\quad\left(\frac{\Phi}{x}\right)_-<\frac{1}{\gamma}.$$

But this result implies that for no value of Φ/x in the allowed range expressed by Eq. (26) does $\Phi'=0$.

Since the viscosity term of Eq. (11′) remains zero so long as $\Phi'>0$, there is no way for the viscosity to enter except by the introduction of discontinuities in the slopes of the physical quantities. This indicates a possible deficiency of the viscosity formalism which in numerical applications assumes continuity not only of the physical quantities themselves but also of their slopes.

IV. FLOW FIELD WITH VISCOSITY

From the above arguments it has been established that Eqs. (13′) to (15′) are solved by a Taylor-type solution in the region $x\geqslant 0$ wherein $G=0$. But in order to introduce the viscosity it has been proved necessary to admit discontinuous slopes. The position at which the discontinuities are located may be prescribed arbitrarily as $x=1$ since this affects only the magnitude of R which is at present arbitrary and fixed only after specifying the constant A. The problem remaining is then the solution for the region $x\geqslant 1$ in which the viscosity is effective. This region comprises the transition from the continuous flow field to the undisturbed

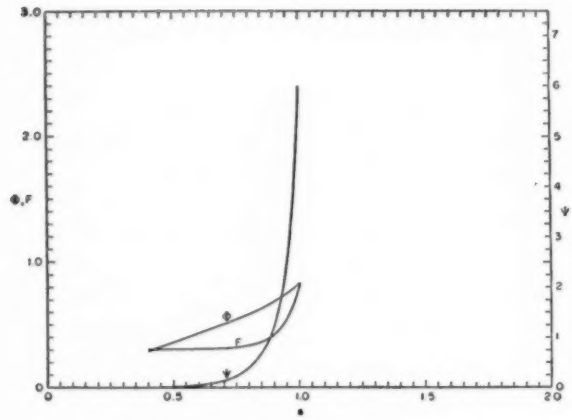


FIG. 1. Solution for $K=0$.

medium in front of the flow. For this region $\Phi' \leqq 0$ and therefore Eq. (11′) may be written

$$\Phi' = -1/Kx(G/\psi)^{\frac{1}{2}}. \tag{11''}$$

At the connection point, $x=1$, $G=0$, as required by continuity of the physical quantities and therefore $\Phi'=0$ at this point. This in turn determines the magnitude of the jump in the slope Φ', which combined with the differential equations prescribes the jumps of the slopes of F, ψ, and G.

The solution of Eqs. (11″) and (13′) to (15′) for the transition region must meet the boundary conditions that for sufficiently large x, the medium is undisturbed so that $\Phi=0$, $F \geqq 0$, $G=0$, and ψ is a constant, and for $x=1$ the solution agrees with the solution for the region $x=0$ to $x=1$, wherein $G=0$. It may be shown quite generally by a series expansion for large x that all solutions of Eqs. (11″) and (13′) to (15′) which extend to infinity behave asymptotically as

$$\Phi \sim \frac{\Phi_0}{x^4}, \quad F \sim \frac{f_0}{x^3}, \quad G \sim \frac{g_0}{x^8}, \quad \text{and} \quad \psi \sim \psi_0 + \frac{\psi_1}{x^5}, \tag{28}$$

where Φ_0, f_0, g_0, ψ_0, and ψ_1 are constants. This result may be verified by direct substitution into the basic equations. This, of course, establishes that the medium in front of the shock transition region is undisturbed and, moreover, is at zero pressure, which implies zero temperature. However, not every solution extends to infinity nor is physically admissible since Φ, F, G, and ψ must be everywhere positive in order to correspond to a diverging flow field with positive pressures and densities. It may be shown that the physically admissible solutions would extend to infinite x. For admissible solutions, the values of the quantities Φ, F, G, and ψ at $x=1$ must be restricted. To examine this restriction, it is observed first that certain restrictions on these quantities already exist. Namely, $G=0$ at $x=1$. Also

$$(F+G)\Phi - (x-\Phi)\left(\frac{F}{\gamma-1}+\frac{1}{2}\psi\Phi^2\right)=0, \tag{29}$$

or at $x=1$

$$F\Phi - (1-\Phi)\left(\frac{F}{\gamma-1}+\frac{1}{2}\psi\Phi^2\right)=0. \tag{30}$$

Equation (29) follows from Eqs. (13′) to (15′) and Eq. (28) by a derivation completely analogous to that leading to Eq. (21); and Eq. (30) is a consequence of $G=0$ at $x=1$. Finally, it is observed from Eqs. (11″) and (13′) to (15′) that all solutions with the same values of F/ψ, G/ψ, and Φ at $x=1$ lead to the same functions $F(x)/\psi(1)$, $G(x)/\psi(1)$, $\psi(x)/\psi(1)$, and $\Phi(x)$. Since $G=0$ at $x=1$ and Eq. (30) relates F/ψ to Φ at $x=1$, the only independent degree of freedom remaining for the choice of the quantities at $x=1$ with which to obtain an admissible solution is the specification of Φ at $x=1$. That, in fact, a choice of Φ can be made which leads to an admissible solution has not been established analytically. However, a convincing demonstration of the existence of such a choice is provided by the explicit numerical solutions discussed in the next section. It may be considered proved, therefore, that Eqs. (11″) and (13′) to (15′) admit of solutions which consist of a continuous hydrodynamic flow from $x=0$ to $x=1$, and which exhibit a continuous transition region (except

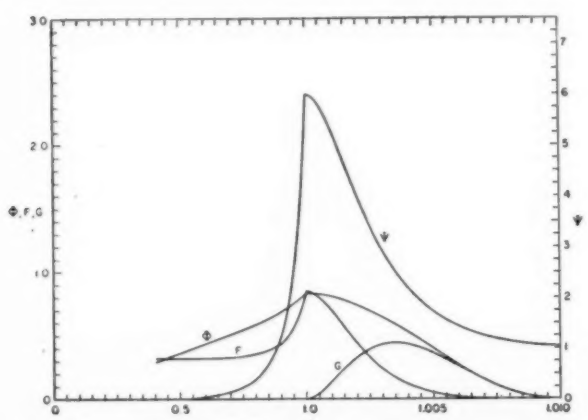


FIG. 2. Solution for $K=0.00349$.

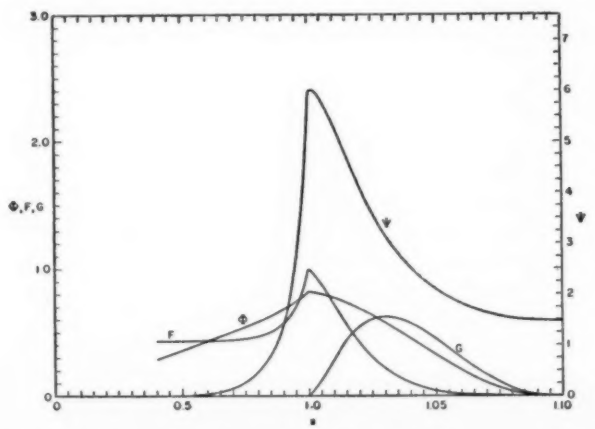


FIG. 3. Solution for $K=0.0349$.

TABLE I. Solutions for various values of K.

$K=0$					$K=0.00349$				
x	Φ	ψ	F		x	Φ	ψ	F	G
0.400	0.287	0.002	0.305		0.400	0.287	0.002	0.318	
0.500	0.358	0.010	0.305		0.500	0.358	0.010	0.319	
0.600	0.430	0.041	0.307		0.600	0.430	0.042	0.320	
0.700	0.503	0.132	0.313		0.700	0.503	0.136	0.327	
0.800	0.584	0.392	0.337		0.800	0.584	0.401	0.351	
0.900	0.685	1.232	0.423		0.900	0.684	1.254	0.439	
0.920	0.710	1.600	0.460		0.920	0.709	1.625	0.476	
0.940	0.737	2.120	0.510		0.940	0.736	2.148	0.527	
0.960	0.766	2.885	0.581		0.960	0.765	2.913	0.599	
0.980	0.799	4.066	0.682		0.980	0.797	4.089	0.700	
1.000	0.833	6.000	0.833		1.000	0.832	6.000	0.852	0.000
					1.002	0.742	3.988	0.441	0.331
					1.004	0.526	2.199	0.121	0.431
					1.006	0.269	1.436	0.021	0.263
					1.008	0.069	1.130	0.001	0.071
					1.010	0.001	1.054	0.000	0.000
$K=0.0349$					$K=0.349$				
x	Φ	ψ	F	G	x	Φ	ψ	F	G
0.400	0.287	0.002	0.430		0.400	0.287	0.004	1.165	
0.500	0.358	0.012	0.431		0.500	0.358	0.020	1.165	
0.600	0.430	0.049	0.433		0.600	0.429	0.077	1.168	
0.700	0.503	0.160	0.441		0.700	0.502	0.248	1.181	
0.800	0.582	0.468	0.469		0.800	0.578	0.705	1.224	
0.900	0.679	1.412	0.569		0.900	0.664	1.928	1.362	
0.920	0.703	1.805	0.610		0.920	0.684	2.375	1.416	
0.940	0.728	2.345	0.665		0.940	0.705	2.946	1.484	
0.960	0.756	3.111	0.741		0.960	0.727	3.686	1.572	
0.980	0.786	4.243	0.847		0.980	0.751	4.666	1.687	
1.000	0.818	6.000	1.000	0.000	1.000	0.777	6.000	1.840	0.000
1.020	0.692	4.130	0.416	0.547	1.200	0.565	7.150	0.475	2.145
1.040	0.451	2.530	0.102	0.571	1.400	0.314	6.256	0.115	1.950
1.060	0.203	1.819	0.014	0.301	1.600	0.123	5.631	0.016	0.985
1.080	0.037	1.532	0.000	0.059	1.800	0.019	5.328	0.001	0.182
1.100	0.000	1.480	0.000	0.000	2.000	0.000	5.275	0.000	0.000

for slope discontinuities) from the flow conditions at $x=1$ to an undisturbed medium at zero pressure for large x.

V. RESULTS AND DISCUSSION

To specify the diverging flow field the constants a and b of Eqs. (20) and (24) [which are equivalent to $\Phi(1)$ and $F(1)/\psi(1)$] and the constant A of Eq. (12) must be fixed. The energy E of Eq. (16′) fixes one relationship among these constants and the specification of the density of the undisturbed medium fixes another relationship. Finally, the requirement of a physically admissible solution in the transition region fixes the last constant. In the limit of vanishing viscosity, $K=0$, the solutions, of course, become identical with the familiar strong-shock flow fields with a jump discontinuity at the shock front, $x=1$.

To illustrate the behavior of the present similarity problem, the solutions of Eqs. (11′) and (13′) to (15′) with $\gamma=1.4$ were obtained from Eqs. (20), (21), and (24) in the region without viscosity and by numerical integration using the Runge-Kutta method in the viscosity region. For $K=0$, that is, no viscosity, the boundary conditions $\Phi(1)=2/\gamma+1$, $F(1)=2/\gamma+1$, and $\psi(1)=\gamma+1/\gamma-1$ define the solution which is shown in Fig. 1. This is just the solution obtained by Taylor. In addition three nonzero values of K were used, namely, $K=0.00349$, 0.0349, and 0.349. Figures 2 to 4 show the solutions for the diverging flow obtained from Eqs. (20), (21), and (24) in the region $0\leq x\leq 1$ and obtained by numerical integration of Eqs. (11″) and (13′) to (15′) in the region $x\geqslant 1$. In Table I, the solutions for these four values of K are tabulated. The values of the integral of Eq. (16′) for these cases are

$$\frac{E}{4\pi\rho_0 A^2}=0.423, \quad \text{for } K=0.00000,$$

$$\frac{E}{4\pi\rho_0 A^2}=0.446, \quad \text{for } K=0.00349,$$

$$\frac{E}{4\pi\rho_0 A^2}=0.642, \quad \text{for } K=0.0349,$$

$$\frac{E}{4\pi\rho_0 A^2}=1.650, \quad \text{for } K=0.349.$$

A comparison of these solutions with the Taylor solution provides an indication of the influence of the viscosity on the flow field. In particular, for the Taylor problem, ρ_0 is the density of the undisturbed medium and $\psi(1)\rho_0$ is the density of the shocked medium. These densities should be compared with $\psi(\infty)\rho_0$ and $\psi(1)\rho_0$

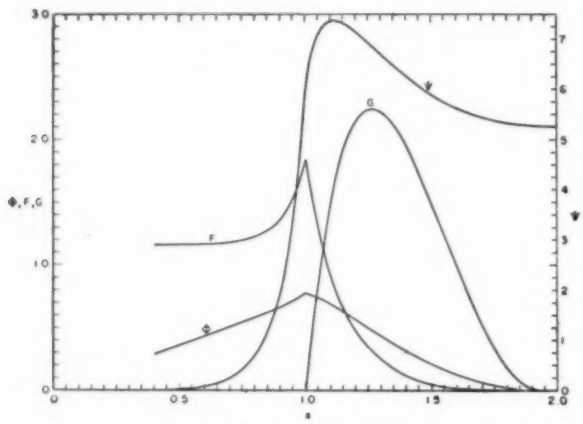


FIG. 4. Solution for $K=0.349$.

of the present problem, which are the densities of the undisturbed and shocked media, respectively. The ratio of these densities is 6 for the Taylor problem, but for $K=0.00349$ it is 5.693, for $K=0.0349$ it is 4.054, and for $K=0.349$ it is 1.137. The use of the viscosity formalism is seen from these ratios, as well as from the complete solutions in Figs. 1 to 4, to introduce considerable modification of the flow field for large K, but negligible changes for small K.

The important consequence of the present discussion is the quantitative estimate of the influence of the viscosity on the flow field. Perhaps as significant is the proof that the viscosity technique may possess the property in spherical flow fields that the physical parameters have discontinuous slopes at the shock transition region, even though these parameters are continuous. Thus the usual numerical integration of Eqs. (1) to (3) and (6), carried out by replacing these equations with difference equations, cannot lead to a solution of the differential equations if the slope discontinuities are disregarded. It would appear therefore that these discontinuities should lead to inaccuracies in numerical applications of the viscosity technique, which ignore them. However, discontinuities of slope exist in other hydrodynamic phenomena, such as rarefaction waves, and apparently lead to no difficulties in numerical treatment. Moreover, from Brode's[3] results on the numerical integration of the hydrodynamic equations (1) to (3) using the viscosity technique for a spherical explosion, it appears that any errors from this source are probably unimportant. This conclusion should be qualified somewhat since in Brode's integration, the radius to the transition region was large and the neighborhood of the shock front was approximately planar so that any discontinuities would be small. Moreover, the form of viscosity used by Brode differed from the present one, although the form which he used also suffered from not leading to a constant mass within the transition region.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to Mr. Jack Little and Miss Ruth Merrill for numerical assistance.

[3] H. Brode (to be published).

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Geometrical Acoustics. II. Diffraction, Reflection, and Refraction of a Weak Spherical or Cylindrical Shock at a Plane Interface

K. O. Friedrichs and Joseph B. Keller
Institute of Mathematical Sciences, New York University, New York 3, New York
(Received November 6, 1954)

The method of a previous paper is applied to the reflection and refraction of a weak spherical or cylindrical shock wave at a plane interface and by a plane slab. The method is also extended to apply to the diffracted wave which appears in these problems whenever total reflection occurs. From the results the leading terms in the asymptotic expansions for high frequency of the fields produced by periodic point or line sources over a plane interface are also obtained. These terms, which apply to electromagnetic as well as acoustic fields, are in exact agreement with results previously obtained by much more complicated methods.

I. INTRODUCTION

We consider a weak spherical or cylindrical shock produced by a point (or line) source (e.g., an explosion) located in the upper half-space $y>0$, which half-space is filled with a homogeneous medium. The lower half-space $y<0$ is filled with another homogeneous medium. When the original shock hits the interface $y=0$ a reflected shock and a refracted shock are produced. The locations and strengths of all these shocks can be easily determined by the geometrical methods derived in "Geometrical Acoustics. I: The Theory of Weak Shocks."[1] This will be done in Sec. II. When the sound speed in the upper medium is greater than or equal to that in the lower medium, no other discontinuities arise and these calculations complete the description of all the shocks in the problem.

However, if the sound speed in the lower medium exceeds that in the upper medium, total reflection occurs at points of the interface where the angle of incidence exceeds the critical angle. This occurrence of total reflection is related to another phenomenon—the appearance of a diffracted shock in the upper medium. The location of this diffracted shock is found in Sec. III by a modification of Fermat's principle. According to the weak shock theory the pressure is continuous across the diffracted shock, but the first time derivative of the pressure with respect to time is discontinuous across it. The jump of this first derivative of the pressure is also determined in Sec. III, and this is the main result of the present paper.

The method devised for the treatment of the diffracted shock is applicable to other problems as well as to the present one. In fact, in Sec. IV, the problem of a spherical shock incident on a layer of one material lying above a half-space containing another material is considered by the same method.

Of course both of these problems can be solved by a different method, that of Fourier decomposition of the source into periodic components.[2] One solves the problem for each periodic component and synthesizes the solution of the shock problem by forming a Fourier integral of the periodic solutions. The result is a Fourier integral of a contour integral of Bessel functions. (Another form of solution can be obtained without recourse to Fourier analysis, using the method of John.)[3] This solution contains the results of the present paper, as well as the pressure distribution behind the shocks. To extract the present results from the solution would require considerable analysis. In particular, such an analysis would require an asymptotic expansion for large ω, of the periodic solution with angular frequency ω. There has been some controversy about this asymptotic expansion,[4] which is avoided by the present method. In fact, in Sec. V, the leading terms in this expansion are obtained from the results of the previous sections, thus avoiding the difficulties often encountered in deriving this expansion.

Summarizing, the advantages of the present method seem to lie in its simplicity, geometric appeal, and generality. On the other hand, while it yields information about the discontinuities or shocks directly, it fails to determine the pressure field behind the shocks.

All the results of the present paper apply equally well to electromagnetic pulses and periodic waves if p is interpreted as the vertical or y component of the Hertz vector and $n=3$ (see reference 2).

II. FORMULATION AND SOLUTION IN THE ABSENCE OF TOTAL REFLECTION

We consider the shock emitted at $t=0$ from a point source of unit strength located at $x=0$, $y=h$ in the homogeneous medium filling the half-space $y>0$. This medium has sound speed c_1 and density ρ_1, while the second medium in the region $y<0$ has sound speed c_2 and density ρ_2. Because no added difficulty arises from doing so, we will simultaneously consider the two-dimensional problem in which the point source is replaced by a line source parallel to the z-axis. The integer n will denote the dimension, which is either 3 or 2.

If r denotes distance from the source then at any time t the incident shock front is given by $r=c_1t$, $y\geq 0$ and the jump in pressure across it, which we denote by p, is

[1] J. B. Keller, J. Appl. Phys. **25**, 938–947 (1954).
[2] E. Gerjuoy, Communs. Pure Appl. Math. **VI**, 73–91 (1953).
[3] F. John, Communs. Pure Appl. Math. **VII**, 245–270 (1954).
[4] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Company, Inc., New York, 1948), pp. 573–583.

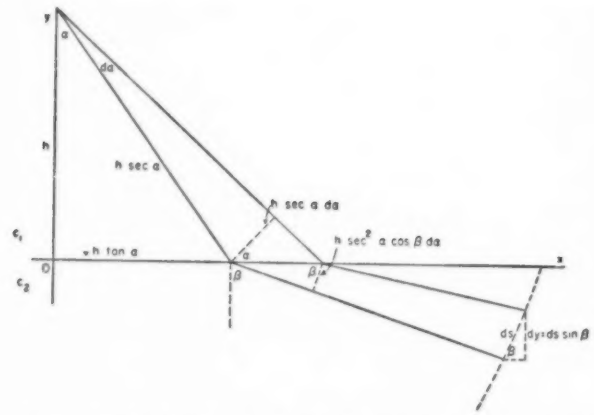


FIG. 1. A pair of neighboring incident rays and the corresponding refracted rays. The angles of incidence are α and $\alpha+d\alpha$, respectively, and the corresponding angles of refraction are β and $\beta+d\beta$. The height of the source is h and the normal separation between the incident rays is $h\sec\alpha d\alpha$ at the interface. The normal separation between the refracted rays is $h\sec^2\alpha\cos\beta d\alpha$ at the interface. From this the normal separation ds between these same rays can be found at any point along these rays.

given by $p=r^{1-n/2}$. These results follow at once from those of the first communication.[1] When $t>hc_1^{-1}$, both reflected and transmitted shocks will be present in addition to the incident shock. The reflected shock is given by the equations $r=c_1t$, $y\geq 0$, $t\geq hc_1^{-1}$, where r' is the distance from the image source $x=0$, $y=-h$.

The strength of the reflected shock is

$$p=(r')^{(1-n)/2}\frac{1-Z}{1+Z},\quad Z=\frac{\rho_1c_1\cos\beta}{\rho_2c_2\cos\alpha}. \tag{1}$$

The result Eq. (1) also follows from, reference 1, the first factor accounting for spreading of the front and the second being a reflection coefficient. The reflection coefficient depends upon the angle of incidence α which the incident ray makes with the normal to the surface $y=0$ (see Fig. 1). Thus p varies from point to point on the reflected shock. The angle β is the angle of refraction, defined by

$$\sin\beta=\frac{c_2}{c_1}\sin\alpha. \tag{2}$$

The equation of the transmitted shock front is most easily given in terms of the angle α by the parametric equations

$$y=-c_2(t-hc_1^{-1}\sec\alpha)\cos\beta,\quad y\leq 0 \tag{3}$$

$$x=c_2(t-hc_1^{-1}\sec\alpha)\sin\beta+h\tan\alpha. \tag{4}$$

These equations are obtained by considering the distance which a shock travels along a ray during the time t (see Fig. 1).

To determine the strength of the transmitted shock, we note that the strength of the incident shock is $(h\sec\alpha)^{(1-n)/2}$ at $y=0$ on the ray α. If we multiply this by the transmission coefficient $2(1+Z)^{-1}$ we obtain the strength of the transmitted shock at $y=0$ on the corresponding transmitted ray. To find the strength of the transmitted shock at any other point on the transmitted ray, we utilize the principle derived in reference 1 that *the shock strength varies inversely as the cross-sectional area of an infinitesimal tube of rays*. From Fig. 1 we see that the cross-sectional area of such a tube is $2\pi(h\tan\alpha)^{n-2}h\sec^2\alpha\cos\beta d\alpha$ at the interface, and $2\pi x^{n-2}(\sin\beta)^{-1}dy$ at a point of the tube with coordinate x. In these considerations, the tube is thought of as bounded by two cones in the three-dimensional case. Combining the preceding results we have for the pressure p at a point on a transmitted ray

$$p=\frac{2}{1+Z}(h\sec\alpha)^{(1-n)/2}\times\left[h\sec^2\alpha\sin\beta\cos\beta(x^{-1}h\tan\alpha)^{n-2}\frac{d\alpha}{dy}\right]^{\frac{1}{2}}. \tag{5}$$

The derivative $dy/d\alpha$ required in Eq. (5) is found, from Eq. (3), to be

$$\frac{dy}{d\alpha}=c_2(t-hc_1^{-1}\sec\alpha)\frac{c_2^2\sin\alpha\cos\alpha}{c_1^2\cos\beta}+\frac{hc_2}{c_1}\cos\beta\sec\alpha\tan\alpha. \tag{6}$$

If we eliminate t from Eq. (6) by means of Eq. (4), when $\beta\neq 0$, and insert Eq. (6) into Eq. (5), we obtain after some simplification

$$p=\frac{2}{1+Z}(\sin\alpha)^{(n-1)/2}\frac{\cos\beta}{\cos\alpha}x^{(2-n)/2}\times\left(x-\left[\frac{c_2^2}{c_1^2}-1\right]h\tan^3\alpha\right)^{-\frac{1}{2}},\quad x-\left[\frac{c_2^2}{c_1^2}-1\right]h\tan^3\alpha>0. \tag{7}$$

The preceding results determine the location and strength of the shocks on all rays for which Eq. (2) defines a real angle β, which is always the case if $c_2/c_1\leq 1$. Thus in this case the analysis is complete.

III. DIFFRACTION

If $c_2/c_1>1$ then for a certain angle α_0, Eq. (2) yields $\beta=\pi/2$. The angle α_0 is determined by

$$\sin\alpha_0=\frac{c_1}{c_2}. \tag{8}$$

For $\alpha>\alpha_0$, Eq. (2) has no real solution for β. In this case, total reflection is said to occur. All of the preceding results apply with the exception of Eq. (1), which does not hold on a totally reflected ray. Instead the pressure

becomes logarithmically infinite at the totally reflected shock front, as has been shown for totally reflected plane shocks.[5]

Even more interesting than the totally reflected shock is the diffracted shock, which appears in the upper medium after total reflection has begun. This shock is shown in Fig. 2. It meets the transmitted shock at the interface $y=0$ and is tangent to the reflected shock at the boundary between its regularly and totally reflected portions. In the two-dimensional case the diffracted shock is a straight line connecting the two points just mentioned, and in three dimensions it is the cone obtained by revolving this line about the vertical line through the source. This shock can be obtained from Huyghens' construction, modified to include circles or spheres of unequal radii centered on the plane $y=0$

Alternatively, the rays which are orthogonal to the diffracted shocks are yielded by a modified form of Fermat's principle, in which the Fermat integral is to be made stationary among all curves connecting a given point to the source and having an arc on the boundary plane $y=0$. The higher sound speed c_2 is assigned to the arc. The resulting ray consists of an incident critical ray (α_0) from the source to the interface, a segment of the corresponding refracted ray which lies in the plane $y=0$, and a final straight line which splits off from the refracted ray at some point on it. The final straight line is in the upper medium ($y\geq 0$), makes the angle α_0 with the normal to the interface $y=0$, and lies in the plane containing this normal and the refracted ray, as well as the incident ray. Thus a single incident ray gives rise to infinitely many diffracted rays, one of which splits off from each point of the refracted ray.[6]

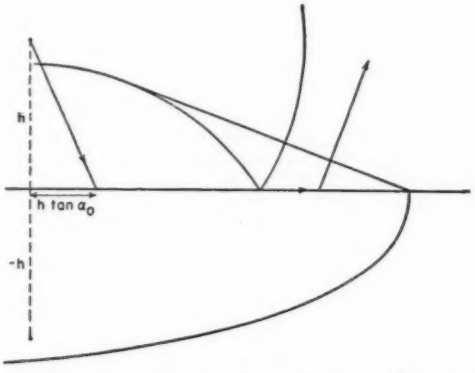


FIG. 2. Incident, reflected, transmitted and diffracted shocks resulting from a spherical shock incident on a plane interface. The sound speed in the lower medium is greater than that in the upper medium. The source is at a height h above the interface, and α_0 is the critical angle. The straight lines with arrows on them are the three portions of a ray. The first part is an incident critical ray, the second is the refracted ray along the interface to which the critical ray gives rise, and the third is a diffracted ray which splits off from the horizontal refracted ray. The diffracted ray is orthogonal to the diffracted shock.

[5] L. Howarth, Proc. Cambridge Phil. Soc. 44, 380-390 (1948).
[6] J. B. Keller, "The geometrical theory of diffraction," Proceedings of the Symposium on Microwave Optics, McGill University, June, 1953.

Since the cross-sectional area of all the incident critical rays is zero, while that of any finite bundle of diffracted rays, to which they give rise, is not zero, it follows that the jump in pressure is zero across the diffracted wave fronts. This can also be seen from Eq. (7), which shows that on the limiting refracted ray $\beta=\pi/2$ which travels along the interface, the pressure p is zero at the shock. However, the first time derivative of the pressure is not continuous, but suffers a finite jump across the diffracted shock, as we will now show. We will also determine the value of the jump in this derivative p_t at all points of the diffracted shock.

To this end we first point out that a jump in p_t varies along a diffracted ray—where p does not jump—in the same manner as does a jump in p, along an ordinary ray, i.e., inversely as the square root of the cross-section of an infinitesimal tube of rays. The proof is similar to that given in reference 1, and therefore we will omit it. Utilizing this fact, we have from Fig. 2,

$$p_t(x,y)=p_t(x-y\tan\alpha_0,\,0)\left[\frac{x-y\tan\alpha_0}{x}\right]^{(n-2)/2};$$

$$x\geq(y+h)\tan\alpha_0,\quad y\geq 0. \tag{9}$$

Here $p_t(x,y)$ denotes the jump of p_t across the diffracted shock when it passes through the point (x,y), and $p_t(x-y\tan\alpha_0,\,0)$ is the corresponding jump on the same ray at the interface $y=0$. Thus the determination of $p_t(x,y)$ is reduced to the determination of p_t at $y=0$.

In order to determine p_t at $y=0$, we note that near $y=0$ the diffracted shock is outside the incident shock; in fact, it is the first shock to arrive. The excess pressure p is therefore zero in front of the diffracted shock and also on it, since the pressure is continuous across it. Thus $p=0$ on the diffracted shock, at least near $y=0$. If we now compute the total time derivative of p at a point on this moving diffracted shock we obtain

$$p_t+c_1\frac{\partial p}{\partial n}=0. \tag{10}$$

Here we have used the fact that the normal velocity of the shock is c_1 and that the tangential derivatives of p are zero: $\partial p/\partial n$ is the derivative of p normal to the shock. Equation (10) permits us to express p_t in terms of $\partial p/\partial n$. Now in the plane $z=0$ we have

$$\frac{\partial p}{\partial n}=p_x\sin\alpha_0+p_y\cos\alpha_0. \tag{11}$$

Also, since $p=0$ on the diffracted wave front

$$0=p_x\cos\alpha_0-p_y\sin\alpha_0. \tag{12}$$

From Eqs. (10), (11), and (12) we have

$$p_t=\frac{-c_1}{\cos\alpha_0}p_y. \tag{13}$$

Now at the interface p_y satisfies the boundary condition

$$\frac{p_y^{(1)}}{\rho_1}=\frac{p_y^{(2)}}{\rho_2}. \tag{14}$$

Thus Eq. (13) becomes

$$p_t=\frac{-\rho_1 c_1}{\rho_2 \cos\alpha_0}p_y^{(2)}. \tag{15}$$

But $p_y^{(2)}$ can be evaluated from Eq. (7) since the refracted shock is normal to the interface $y=0$. Thus from Eqs. (7), (15), (3), and (4) we obtain p_t at the interface $y=0$

$$p_t(x,0)=\frac{2\rho_1 c_1(\sin\alpha_0)^{(n-1)/2}}{\rho_2\cos^2\alpha_0}\times(x-h\tan\alpha_0)^{-\frac{1}{2}}x^{(2-n)/2},\quad x-h\tan\alpha_0>0. \tag{16}$$

To find p_t off the interface we combine Eqs. (16) and (9) to obtain

$$p_t(x,y)=\frac{2\rho_1 c_1}{\rho_2}\frac{(\sin\alpha_0)^{(n-1)/2}}{\cos^2\alpha_0}(x-[y+h]\tan\alpha_0)^{-\frac{1}{2}}x^{(2-n)/2},\quad x-(y+h)\tan\alpha_0>0. \tag{17}$$

Equation (17) is the main result of the present analysis since it yields the value of p_t at the diffracted wave front. The quantity $(x-[y+h]\tan\alpha_0)$ is the length of that portion of the ray from the source to x, y which lies in the interface, and according to Eq. (17), $p_t(x,y)$ is inversely proportional to the three-halves power of that length. Therefore on rays for which that length is zero, p_t is infinite. These are just the critical totally reflected rays, and this result is related to the singular behavior of p on such rays which was discussed after Eq. (8). From Eq. (17) we also see that $p_t(x,y)$ is inversely proportional to $x^{(n-2)/2}$, where x is horizontal distance from the source. Thus for fixed y and large x, $p_t(x,y)$ behaves like $x^{-(n+1)/2}$. For $n=3$ we have $p_t\sim x^{-2}$ which is a much more rapid decay than that of p or p_t in a spherical wave, both of which decay like x^{-1}.

IV. A POINT SOURCE ABOVE A LAYER

We will now extend the previous calculations to the case in which medium 2 is of finite thickness H and is bounded below by a third medium of density ρ_3 and sound speed c_3 with $c_3\leq c_2$. In this case all of the previous results apply to the incident, reflected, transmitted, and diffracted waves in both the cases $c_2\leq c_1$ and $c_2>c_1$. However, when the transmitted shock in medium 2 reaches the 2–3 interface, a new reflected shock is produced in medium 2 and a transmitted shock in medium 3. The reflected shock in medium 2 travels upward to the 1–2 interface where a new reflected wave in medium 2 and a transmitted wave in medium 1 are produced. This process continues indefinitely, but no new diffracted waves can appear due to the condition $c_3\leq c_2$ which we have imposed. We will now determine the equations of all these shocks.

To this end, let us characterize a particular shock by a non-negative integer m which equals the number of intersections made with the 2–3 interface by a ray from the source to the shock. By continuity this number is the same for all points on a given shock (i.e., for all rays to that shock). Then by considering the distance which a shock front can travel along a given ray in time t, we obtain the following parametric equations for the various shocks:

Parametric Equations of Shocks

Medium 1

a. Incident shock

$y=h-c_1 t\cos\alpha,\quad y\geq 0$
$x=(h-y)\tan\alpha.$

b. Reflected shocks

$y=c_1 t\cos\alpha-h-2mHc_1c_2^{-1}\cos\alpha\sec\beta,\quad y\geq 0$
$x=(h+y)\tan\alpha+2mH\tan\beta.$

Medium 2

a. Downgoing shocks

$y=-c_2 t\cos\beta+hc_2c_1^{-1}\sec\alpha\cos\beta+2mH,\quad 0\geq y\geq -H$
$x=h\tan\alpha+(2mH-y)\tan\beta.$

b. Upgoing shocks

$y=c_2 t\cos\beta-hc_2c_1^{-1}\sec\alpha\cos\beta-2mH,\quad 0\geq y\geq -H,\ m>0$
$x=h\tan\alpha+(2mH-y)\tan\beta.$

Medium 3

a. Downgoing shocks

$y=-H-c_3 t\cos\gamma+hc_3c_1^{-1}\cos\gamma\sec\alpha+(2m+1)Hc_3c_2^{-1}\cos\gamma\sec\beta,\quad y\leq -H,\ m>0$
$x=h\tan\alpha+(2m+1)H\tan\beta-(y+H)\tan\gamma.$

In medium 1 the reflected shock with $m=0$ is the shock reflected from the 1–2 interface and considered in Sec. II. Similarly the downgoing shock in medium 2 with $m=0$ is the transmitted shock of Sec. II. In medium 3 the angle of refraction γ is defined by

$$\sin\gamma=\frac{c_3}{c_2}\sin\beta. \tag{18}$$

In order to determine the strengths of these shocks, we introduce the following reflection coefficients R_{ij} and transmission coefficients T_{ij}

$$R_{ij}=\frac{1-Z_{ij}}{1+Z_{ij}},\quad T_{ij}=\frac{2}{1+Z_{ij}}. \tag{19}$$

The impedance Z_{ij} is given by

$$Z_{12}=Z_{21}^{-1}=\frac{\rho_1 c_1\cos\beta}{\rho_2 c_2\cos\alpha},\quad Z_{23}=Z_{32}^{-1}=\frac{\rho_2 c_2\cos\gamma}{\rho_3 c_3\cos\beta}. \tag{20}$$

By examining the number of reflections and transmissions of a shock at each interface, we find that the following reflection and transmission factors F must be used in determining the strengths of the various shocks:

Reflection and Transmission Factor F

Medium 1

a. Incident shock:	$F=1.$	
b. Reflected shocks	$F=R_{12}$	$m=0$
	$F=T_{12}R_{23}{}^{m}R_{21}{}^{m-1}T_{21}$	$m>0.$

Medium 2

a. Downgoing shocks	$F=T_{12}R_{23}{}^{m}R_{21}{}^{m}$	$m\geq 0.$
b. Upgoing shocks	$F=T_{12}R_{23}{}^{m}R_{21}{}^{m-1}$	$m>0.$

Medium 3

a. Downgoing shocks	$F=T_{12}R_{23}{}^{m-1}R_{12}{}^{m-1}T_{23}$	$m>0.$

To complete the determination of the shock strengths we must compute, for each shock, the area ratio dA_0/dA. Here dA designates an element of area on the shock under consideration, and dA_0 is the corresponding area on the incident shock at unit distance from the source. Then since $p=1$ on the incident shock at unit distance from the source, we may compute p on any other shock by the energy principle[1]

$$p=F\left(\frac{\cos\delta}{\cos\alpha}\frac{dA_0}{dA}\right)^{\frac{1}{2}}. \tag{21}$$

In Eq. (21) δ is either α, β, or γ according as p is computed at a point in medium 1, 2, or 3. The factor $\cos\delta/\cos\alpha$ is included in Eq. (21) to account for the change in area of a shock front, or tube of rays, when refracted at an interface. The factor F is to be taken from the preceding table for the case under consideration.

To compute dA_0 we consider a pair of rays (or of cones if $n=3$) separated by the angle $d\alpha$. The area of the shock bounded by these rays (cones) at unit distance from the source is

$$dA_0=x_0{}^{n-2}d\alpha=(\sin\alpha)^{n-2}d\alpha. \tag{22}$$

A factor of 2π is to be included if $n=3$, but since dA will then contain a similar factor, it will cancel out of Eq. (21).

To compute dA we note that the angle between the vertical and a ray normal to the shock is just δ defined above. Therefore if ds denotes an element of arc length along the shock in the xy plane, we have $ds=dy/\sin\delta$. Thus we have for dA

$$dA=x^{n-2}ds=x^{n-2}\frac{dy}{\sin\delta}. \tag{23}$$

Using Eqs. (22), (23) in Eq. (21) yields

$$p=F\left[\frac{\cos\delta\,\sin\delta(\sin\alpha)^{n-2}}{x^{n-2}\cos\alpha}\left(\frac{dy}{d\alpha}\right)^{-1}\right]^{\frac{1}{2}}. \tag{24}$$

In order to calculate p from Eq. (24), the absolute value of $dy/d\alpha$ must be employed since this factor arose from area considerations. This derivative can be computed from the shock equation listed in the table above, for the shock under consideration. Carrying out this differentiation, substitution into Eq. (24) and simplifying finally gives for p the following expressions:

Shock Pressure p

Medium 1

a. Incident shock

$$p=(c_1t)^{(1-n)/2}.$$

b. Reflected shocks

$$p=R_{12}(c_1t)^{(1-n)/2},\quad m=0$$

$$p=T_{12}R_{23}{}^{m}R_{21}{}^{m-1}T_{21}\left(\frac{\sin\alpha}{x}\right)^{(n-2)/2}\sin^{-\frac{1}{2}}\alpha$$

$$\times(x-[c_2{}^2c_1{}^{-2}-1]2mH\tan\beta\tan^2\alpha)^{-\frac{1}{2}},\quad m>0.$$

Medium 2

a. Downgoing shocks

$$p=T_{12}R_{23}{}^{m}R_{21}{}^{m}\left(\frac{\sin\alpha}{x}\right)^{(n-2)/2}\frac{\tan\alpha\cos\beta}{\sin^{\frac{1}{2}}\alpha}$$

$$\times(x-[c_2{}^2c_1{}^{-2}-1]h\tan^3\alpha)^{-\frac{1}{2}},\quad m\geq 0.$$

b. Upgoing shocks

$$p=T_{12}R_{23}{}^{m}R_{21}{}^{m-1}\left(\frac{\sin\alpha}{x}\right)^{(n-2)/2}\frac{\tan\alpha\cos\beta}{\sin^{\frac{1}{2}}\alpha}$$

$$\times(x-[c_2{}^2c_1{}^{-2}-1]h\tan^3\alpha)^{-\frac{1}{2}},\quad m>0.$$

Medium 3

a. Downgoing shocks (assuming $c_3=c_1$ which implies $\gamma=\alpha$)

$$p=T_{12}R_{23}{}^{m-1}R_{21}{}^{m-1}T_{23}\left(\frac{\sin\alpha}{x}\right)^{(n-2)/2}\sin^{\frac{1}{2}}\alpha$$

$$\times(x-[c_2{}^2c_1{}^{-2}-1](2m+1)H\tan\beta\tan^2\alpha)^{-\frac{1}{2}},\ m>0.$$

V. ASYMPTOTIC EXPANSION OF PERIODIC WAVES

Suppose the impulsive source considered previously is replaced by a periodic source with the same maximum amplitude and the time dependence given be $e^{-i\omega t}$. Then we may suppose that the entire pressure field has this same time dependence. We now seek an asymptotic expansion of this field for large ω. From the work of Luneburg[7] and Kline[8] the leading term in this expansion—the so-called geometrical acoustics field—can

[7] R. K. Luneburg, "The mathematical theory of optics," Brown University, 1944; and "Propagation of electromagnetic waves," New York University, 1948.

[8] M. Kline, Communs. on Pure Appl. Math. **IV**, 225–263,1951; J. Rational Mech. Analy. **3** No. 3, (May, 1954).

be obtained from the discontinuities in the transient solution previously considered. The next term in this expansion can be determined from the discontinuities in the first time derivative of the transient solution, and further terms can be found in terms of discontinuities in higher time derivatives.

Since we have already investigated all the discontinuity surfaces in the transient problem, we can immediately determine the first two terms in the asymptotic expansion of the periodic field. According to the results of Luneburg and of Kline the leading term in the expansion at any point is a sum of terms, one corresponding to each discontinuity in the pulse solution at that point. In the problem considered in Secs. II and III, in which only two media are present, there are two discontinuities in p at a point in medium 1 and only one discontinuity in medium 2. The discontinuities in medium 1 are the incident and reflected shocks; that in medium 2 is the transmitted shock. Therefore we have

$$\begin{aligned} p_\omega &\sim [p_I]e^{i(\omega t_I-\omega t)}+[p_R]e^{i(\omega t_R-\omega t)} && \text{in medium 1,} \\ p_\omega &\sim [p_T]e^{i(\omega t_T-\omega t)} && \text{in medium 2.} \end{aligned} \tag{25}$$

Here $[p]$ denotes the jump in p in the transient problem and the subscripts I, R, and T refer, respectively, to the incident, reflected, and transmitted shocks. The time at which any one of these shocks passes the point at which p_ω is being evaluated is denoted by t_I, t_R or t_T. The quantity p_ω is the pressure in the periodic field of angular frequency ω, and the subscript is employed to distinguish it from p, the pressure in the transient problem. The terms shown in Eq. (25) yield what is called the geometrical acoustics field, and can be written more explicitly by using Eqs. (1), (4), and (7). Then Eq. (25) becomes (on regularly reflected rays)

$$\begin{aligned} p_\omega \sim{}& \frac{\exp\left[i\left(\frac{\omega}{c_1}r-\omega t\right)\right]}{r^{(n-1)/2}}+\frac{1-Z}{1+Z} \\ &\times\frac{\exp\left\{i\left(\frac{\omega}{c_1}r'-\omega t\right)\right\}}{r'^{(n-1)/2}} \quad \text{in medium 1,} \\ p_\omega \sim{}& \frac{2}{1+Z}(\sin\alpha)^{(n-1)/2}\frac{\cos\beta}{\cos\alpha}x^{(2-n)/2} \\ &\times\left(x-\left[\frac{c_2^2}{c_1^2}-1\right]h\tan^2\alpha\right)^{-\frac{1}{2}} \\ &\times\exp\left\{i\left(\frac{\omega}{c_1}\left[\frac{x-h\tan\alpha}{c_2\sin\beta}+hc_1^{-1}\sec\alpha\right]-\omega t\right)\right\}; \\ &\left(x-\left[\frac{c_2^2}{c_1^2}-1\right]h\tan^2\alpha\right)>0. \end{aligned} \tag{26}$$

In the three medium problem of Sec. IV, there are infinitely many shocks passing through each point and the corresponding geometrical acoustics field contains infinitely many terms.

The foregoing results are, of course, well known and have been included here merely for the sake of completeness and also to exemplify the Luneburg-Kline theory. The main point of the present paper was the determination of the discontinuity in p_t across the diffracted shock when it occurs. The knowledge of this discontinuity enables us to obtain another term, which is of order $1/\omega$, in the asymptotic expansion in medium 1. This term is

$$-\frac{1}{i\omega}[p_t]_D\exp\left[i\left(\frac{\omega}{c_1}t_D-\omega t\right)\right]. \tag{27}$$

Here the subscript D denotes that the quantity bearing the subscript refers to the diffracted shock. Therefore this term appears only in that part of medium 1 through which the diffracted shock passes. According to Eq. (17) this term becomes

$$\begin{aligned} &\frac{-2\rho_1c_1}{i\omega\rho_2}\frac{(\sin\alpha_0)^{(n-1)/2}}{\cos^2\alpha_0}(x-[y+h]\tan\alpha_0)^{-\frac{1}{2}}x^{(2-n)/2} \\ &\times\exp\left\{i\left(\frac{\omega}{c_1}[(y+h)(1-\alpha_0^2)^{\frac{1}{2}}+\alpha_0x]-\omega t\right)\right\}. \end{aligned} \tag{28}$$

This result, for three dimensions, as well as Eq. (26), are identical with the results of Gerjuoy,[2] who obtained them by a completely different and much more complicated method. He solved the time-periodic problem exactly by separation of variables, and then evaluated the resulting integral form of the solution asymptotically. In his paper the solution is discussed, the applicability of the result to certain electromagnetic fields is pointed out, and numerous references to other papers dealing with this problem are given.

There is, however, another term of order $1/\omega$ in the asymptotic expansion in medium 1, and a term of this same order in medium 2. These terms correspond to the discontinuity in p_t at the reflected shock in medium 1 and at the transmitted shock in medium 2. They were not obtained by Gerjuoy nor, presumably, by other authors, although they could be deduced from the exact solution by a more precise asymptotic evaluation of the integral. They will not be determined here either, but will be given in a forthcoming article devoted to asymptotic expansions.[9]

[9] Keller, Lewis, and Seckler, Asymptotic Solution of Some Diffraction Problems, Inst. Math. Sciences, New York University, Research Report EM-80, June, 1955.

X-Ray Path through Calcite in Laue Diffraction*

GEORGE L. ROGOSA AND GUENTER SCHWARZ
Department of Physics, Florida State University, Tallahassee, Florida
(Received February 14, 1955)

The position and width of the x-ray beam transmitted through calcite in Laue diffraction has been measured. For the extinction case, the transmitted beam has the same position and width as the incident beam. For the anomalous transmission case, the transmitted beam widens showing that the radiation propagates through the crystal along the diffracting planes. These results are in general accord with the theoretical calculations of von Laue.

I. INTRODUCTION

THE intensity of the radiation transmitted through a single crystal in Laue diffraction may increase, decrease, or go through a maximum and minimum as the crystal is rotated through the diffracting position. This phenomenon was first found by Borrmann[1] and has since been extensively investigated by others.[2–4] The case of decreased transmitted intensity is referred to as extinction, whereas increased transmitted intensity is referred to as anomalous transmission. At a wavelength of 0.708 A a crystal of calcite having a thickness of 0.4 mm belongs in the extinction class while crystals with a thickness above 2 mm belong in the anomalous transmission class.

von Laue[5] has shown theoretically that for the case of anomalous transmission, the radiation is propagated through the crystal as a single wave field mainly along the diffracting planes and upon leaving the crystal splits up into a reflected and a transmitted wave.

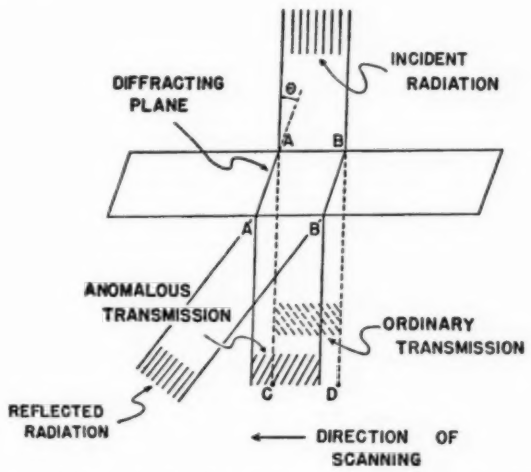


FIG. 1. Schematic drawing of the radiation path in the crystal. The angles in the drawing are not to scale in order to clearly illustrate the results.

Figure 1 is a schematic drawing of the path of the radiation as it travels through the crystal. For ordinary transmission, the emerging beam is shown as CD. For anomalous transmission the beam would emerge at A′B′ and continue as shown. We have measured the position of the transmitted x-ray beam with respect to the incident beam in the case of extinction and anomalous transmission.

II. EXPERIMENTAL

Figure 2 shows the general experimental arrangement.[6] A beam of Mo $K\alpha_1$ radiation coming from a double crystal monochromator is limited to a width of about 0.6 mm by a slit consisting of two lead blocks and then passes over the axis of a spectrometer on which the calcite crystal under investigation is mounted. The transmitted radiation enters a 70 cm pressure krypton Geiger counter connected to a counting rate meter and graphic recorder. A 0.06 mm wide gold slit is positioned in front of the counter window. The counter together with the slit in front is driven across the transmitted beam in the direction shown in Fig. 1 with the crystal at the diffracting position and then

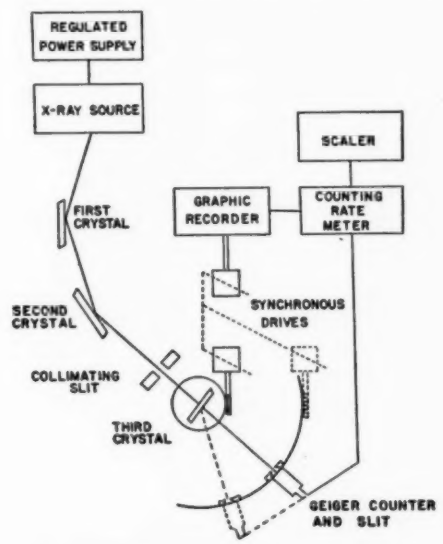


FIG. 2. General experimental arrangement.

* This work was supported by a grant from the National Science Foundation.

[1] G. Borrmann, Physik. Z. **43**, 157 (1941); Z. Physik **127**, 297 (1950).
[2] H. N. Campbell, J. Appl. Phys. **22**, 1139 (1951).
[3] G. Schwarz and G. L. Rogosa, Phys. Rev. **95**, 950 (1954).
[4] G. Brogen and O. Odell, Arkiv Fysik **8**, 401 (1954).
[5] M. von Laue, Acta Cryst. **5**, 619 (1952).

[6] For a more complete description of the equipment in this experimental set up, see reference 3.

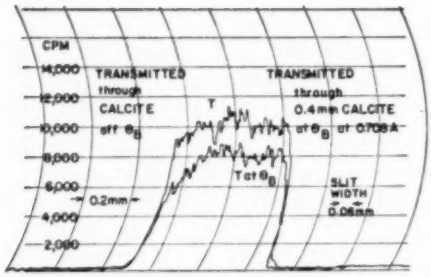


FIG. 3. Cross section of transmitted beams for a 0.4 mm crystal at and off the diffracting position.

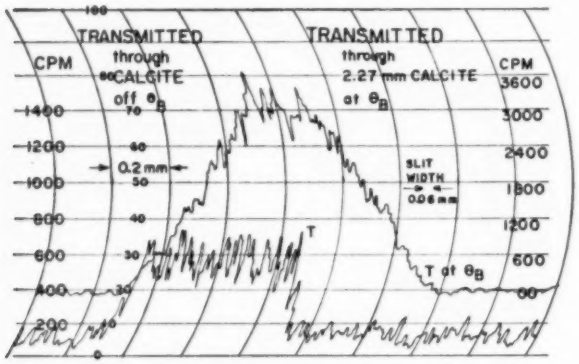


FIG. 4. Cross section of transmitted beams for a 2.27 mm crystal at and off the diffracting position.

again with the crystal off the diffracting position or removed altogether.

III. RESULTS AND DISCUSSION

Figure 3 shows the results for a 0.4 mm thick calcite crystal in the extinction case. The position and width of the transmitted beam remained fixed, within the limits of detection, for the crystal at and off the diffracting position. Figures 4 and 5 show the results for the case of anomalous transmission with 2.27 mm and 3.5 mm calcite crystals. Edge D of the transmitted beam stays fixed while the other edge undergoes a displacement in the direction of A′ for which radiation propagated along the crystal planes would be found.

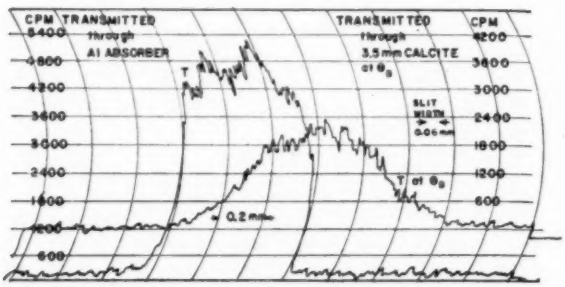


FIG. 5. Cross section of transmitted beams for a 3.5 mm crystal at the diffracting position and with the crystal removed.

Radiation which travels entirely along the crystal planes will be displaced, in this arrangement, 12% of the crystal thickness. For the 3.5 mm crystal this calculated displacement is 0.42 mm whereas Fig. 5 shows a total displacement of 0.6 mm. This extra displacement, somewhat weaker in intensity, can be qualitatively understood by assuming that not all the radiation follows the crystal planes through the entire crystal. Some radiation may be internally diffracted and cross the crystal planes for some distance within the crystal, shifting it beyond A′. This radiation while crossing the planes would have the normal absorption while that which follows the planes would have the reduced absorption. Radiation which continues into the crystal in the straightforward direction and is diffracted at some depth in the crystal, then following the crystal planes, does not undergo the full 12% displacement as seen in Figs. 4 and 5. In fact, the coincidence of the transmitted beam with the incident beam at position D indicates that some radiation goes through the crystal undeviated in path. The results given here are in agreement with the earlier proposals of Borrmann and the theoretical considerations of von Laue.

ACKNOWLEDGMENTS

We acknowledge the valuable assistance of Mr. L. Carlton Brown in working out the experimental arrangement.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Weak Detonations and Condensation Shocks

W. H. Heybey and S. G. Reed, Jr.*
U. S. Naval Ordnance Laboratory, White Oak, Maryland
(Received November 6, 1954)

The absence of weak detonations in available experience is felt to pose a problem since the well-known rapid condensation processes encountered in the flow of expanding gases have all the characteristics and present, in fact, examples of weak detonations. As such, they would suggest that the initiation mechanism producing weak detonations will have to be different from that commonly adopted for strong detonations. The physical aspects of rapid condensation processes are discussed and compared with the conditions obtained by Friedrichs for the existence of weak detonations.

INTRODUCTION

RAPID condensation of gas constituents in supersonic nozzle flow leads to shock-like phenomena which have been known for some time[1,2] and have been discussed by various authors.[3] The exothermic character of these processes involving phase changes places them in close relation to flow processes involving chemical reactions. This relationship has been stated speculatively by J. M. Burgers[4] and has been touched on more recently by K. Oswatitisch.[5] In the present paper a more detailed treatment is presented by describing the place these condensation processes occupy within the general framework of flow discontinuities induced by exothermic processes. Such a place exists because the laws of conservation of mass, momentum, and energy are the foundation of the treatment of all these flow discontinuities. It then becomes apparent that irreversible condensation processes in supersonic flow are formally identical with weak detonations and may be viewed as examples of such. We call them here "weak condensation detonations."

Chemical weak detonation fronts have never been definitely observed.[6] Theoretically, this gap in available experience is commonly linked to the initiation process.[7,8] However, the possible existence of three-dimensional weak detonations has been demonstrated by Taylor[9] and of one-dimensional detonations by Friedrichs.[10] Taylor gives no hint as to the physical conditions under which his solution might be realized; Friedrichs, however, does link the existence problem with the behavior of chemical reaction rates.

In the following we present, first, a short discussion of the normal condensation shocks in a steady supersonic air stream containing a small portion of water vapor. This enables a comparison to be made with the standard treatment of steady detonation theory. Next, the case of rapid condensation of a pure substance, e.g., steam, in steady nozzle flow is briefly dealt with. Both of these examples show that these condensation processes are, in fact, weak detonations. A discussion of the kinematical features of such a process in one-dimensional unsteady flow is then given and compared with shock tube experiments. Finally, the physical aspects of condensation processes are discussed and related to the conditions for chemical weak detonations issuing from the theoretical work of Freidrichs.

STEADY CONDENSATION SHOCKS IN AIR INDUCED BY SUDDEN CONDENSATION OF A SMALL FRACTION OF WATER VAPOR

Shock phenomena in wind tunnel nozzles ascribed to the sudden condensation of a small amount of water have often been observed before air drier installations came into use.[1,2] The simple theory[11] of these transitions considers the condensation merely as a heat source and deals with the changes wrought upon a uniform air stream by steady injection of heat energy, say of g calories per unit mass, at a fixed station. This addition instantaneously raises the original stagnation enthalpy (from h_0 to h_0+g calories per unit mass) and, consequently, the sound speed at a fictitious throat from a^* to $a^{*\prime}$. The discontinuous change of state is governed by the three conservation laws (of mass, momentum, and energy) which permit a complete description of the transition. In what follows the states immediately in front of and behind the (normal) shock will be designated by the subscripts 1 and 2, respectively. The dimensionless flow speeds $v_1=u_1/a^*$ and $v_2=u_2/a^{*\prime}$ are related by the conservation laws which yield

$$2Qv_1v_2=(1+v_1^2)\pm[(1+v_1^2)^2-4Qv_1^2]^{\frac{1}{2}}, \qquad (1)$$

* Now at the Office of Naval Research.

[1] C. Wieselsberger and L. Prandtl, Proceedings of the Volta Conference, Rome (1935); C. Robb, Can. J. Research **A19**, 87, 167 (1941).

[2] R. Hermann, Luftfahrt-Forsch. **19**, 201 (1942).

[3] A review of background on condensation shocks and further references can be found in the article by P. Wegener (see reference 13).

[4] J. M. Burgers, address at the dedication of the Naval Ordnance Laboratory, Aeroballistic Research Facilities, NOLR 1130 (1949).

[5] K. Oswatitisch, *Gasdynamik* (Springer-Verlag, Wien, 1951), pp. 59, 65.

[6] A. R. Ubbelohde, Proc. Roy. Soc. (London) **A204** (1950); Proc. IV, *International Conference on Combustion and Detonation* (Williams and Wilkins, Baltimore, Maryland, 1952).

[7] R. Courant and K. O. Friedrichs, *Supersonic Flow and Shock Waves* (Interscience Publishers, Inc., New York, 1948).

[8] J. Von Neumann, Office of Scientific Research and Development, Rept. 549 (1942).

[9] G. I. Taylor, Proc. Roy. Soc. (London) **A200**, 235 (1950).

[10] K. O. Friedrichs, NAVORD Report 79-46.

[11] W. Heybey, National Advisory Council of Aeronautics, TN 1174 (1947).

where

$$Q^2=1+g/h_0=(a^{*\prime}/a^*)^2.$$

It is seen that the velocity v_2 is not determined by v_1 alone; the quantity g must also be known which cannot be determined from gas dynamical relations alone and will be taken here as an indeterminate parameter. In order for v_2 to be real, however, g obviously cannot be larger than

$$g_{\max}=h_0\frac{(v_1^2-1)^2}{4v_1^2}. \tag{2}$$

This condition is imposed by the three conservation laws.

This ambiguity in sign appearing in Eq. (1) can be removed. Since normal condensation shocks have been observed in supersonic flow only, we first take it that $v_1>1$. Secondly, no change of state can occur in the uniform air stream if no liquid is formed at all, so that $v_2=v_1$ if $Q=1$. These conditions require the upper sign in Eq. (1). Once the velocity v_2 is thus established as a unique function of v_1 and g, a simple analysis, omitted here, shows that a normal condensation shock transforms supersonic velocity at the front side into a (smaller) supersonic velocity at the back side, at the same time raising pressure and density. It thus provides for the same change of state as would a one-dimensional weak detonation front separating unburnt gas from burnt gas in a chemical reaction process.[7] The part of the air stream containing liquid water obviously corresponds to the burnt gas flow.

RELATIONSHIP WITH WEAK DETONATIONS

For a mathematical treatment in terms of a Hugoniot diagram, one will observe that the internal energy, e, per unit mass of air, assumed to be a perfect gas, is the same function of specific volume, τ, and pressure, p, before and behind the front:

$$e=\frac{1}{\gamma-1}p\tau.$$

The equation of the Hugoniot curves then becomes

$$\frac{\gamma+1}{\gamma-1}(p_1\tau_1-p\tau)+\tau_1 p-p_1\tau+2g=0. \tag{3}$$

It is seen that for parametric values of g the curves here are nonintersecting hyperbolas with common asymptotic lines

$$p=-\frac{\gamma-1}{\gamma+1}p_1,\quad \tau=\frac{\gamma-1}{\gamma+1}\tau_1.$$

Adopting now the pattern used in conventional theory to exclude strong deflagrations and weak detonations,[7,8] we replace the ideally sharp shock discontinuity by a narrow reaction zone within which the heat is liberated gradually. In other words, the agglomeration of water molecules into droplets is not supposed to occur instantaneously, since it requires a small but finite time to be completed. It thus extends across a narrow slab, which in turn is subdivided into individual layers, each characterized by the amount of heat liberated up to that layer. The Hugoniot curve, Eq. (3), with $g=g_i$ gives the possible $p-\tau$-relationship in that layer; the values actually prevailing are obtained by the intersection with the Rayleigh line,

$$\frac{p-p_1}{\tau-\tau_1}=-\rho_1^2u_1^2=-\gamma\frac{p_1}{\tau_1}M_1^2$$

where ρ_1 denotes the density forward of the shock and M_1 is the supersonic Mach number at which condensation sets in.

The slope of the Hugoniot hyperbola at P_1, according to Eq. (3), is $-\gamma p_1/\tau_1$. Since $M_1>1$, it follows that at P_1 the inclination of the Rayleigh line is steeper than that of the Hugoniot curve. Hence, the intersection points on Fig. 1 will, in chemical reaction processes, correspond to detonations. We denote by g_f the final amount of heat liberated by the condensation process, thus assigning a definite value to the parameter g. It is geometrically clear that g_f cannot be larger (for a given Rayleigh line) than a certain amount $g_{\max}$ [which is given analytically by Eq. (2)]. The Hugoniot hyperbola $g=g_{\max}$ will have a point of contact with the Rayleigh line on Fig. 1. With detonations this point describes a Chapman-Jouguet process[7,8,12] for which the velocity is sonic in the burnt gas, and indeed behind a condensation shock the velocity u_2 is known to be sonic if the maximum amount of heat is liberated.[8]

ESSENTIAL FEATURES OF WEAK CONDENSATION DETONATIONS

The successive states within the narrow reaction zone are given by a sequence of points on the Rayleigh line. This sequence starts out at P_1 and comprises all

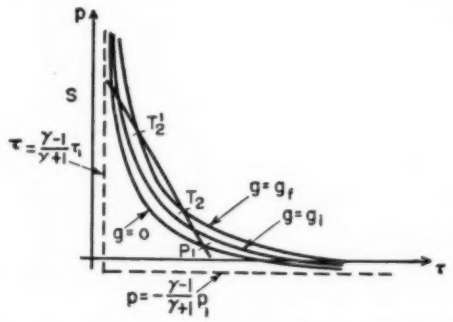


FIG. 1. Hugoniot curves and Rayleigh line for condensation of a fraction of water vapor.

[12] S. Brinkley and J. M. Richardson, Proc. IV, *International Symposium on Combustion* (Williams and Wilkins, Baltimore, 1954), p. 450.

the intersection points with the Hugoniot curves for $g=g_i$, until the point T_2 on the hyperbola $g=g_f$ is reached. With chemical reactions all these points would characterize the momentary state within a weak detonation zone. Beyond T_2 the sequence cannot be continued, because this would call for layers where more than the final amount of heat is liberated. In particular, the strong detonation point T_2' cannot be reached. This reasoning evidently is parallel to that commonly used in excluding strong deflagrations[7] and, therefore, weak detonations, since these are visualized as initiated by a shock setting off a strong deflagration. It appears that this mechanism cannot be upheld for weak condensation detonations; since they are real, they cannot begin with a shock. They may have a shock-like appearance, however, if the phase change occurs over a sufficiently short distance. Measurements by Wegener[13] reveal that the reaction zone, as indicated by nonisentropic gradual pressure variation, can be quite wide (order of magnitude 1 cm). It should also be emphasized that Friedrichs,[10] in establishing the existence of weak detonations, never makes use of a shock transition in the unburnt gas.

The Chapman-Jouguet process appears as a limiting case of weak detonations. One can adapt the arguments usually given[7] for the "stability" of the Chapman-Jouguet point when approached from the "strong" side along a fixed Rayleigh line to show that this point has also an analogous stability when approached from the weak detonation side. A weak detonation cannot proceed past the Chapman-Jouguet point.

CASE OF SUDDEN CONDENSATION OF A SINGLE COMPONENT GAS

There is, of course, only one Hugoniot curve for the final equilibrium state that can be attained in a steady, adiabatic process starting from a given initial state. This can be shown for the case of the mixture of water vapor and air discussed in the foregoing; for simplicity we give here the Hugoniot diagram for pure water vapor. Using thermodynamic data for steam, Hugoniot diagrams can be drawn for equilibrium conditions downstream of any given point. Figure 2 illustrates the

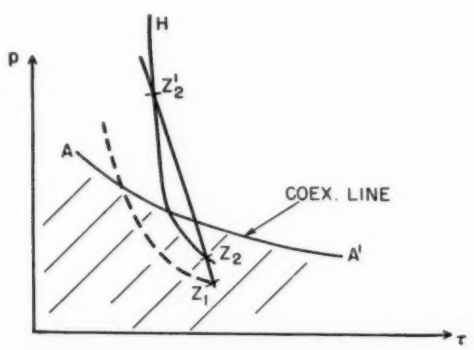


FIG. 2. Hugoniot diagram for steam.

[13] P. Wegener, J. Appl. Phys. **25**, 1485 (1954).

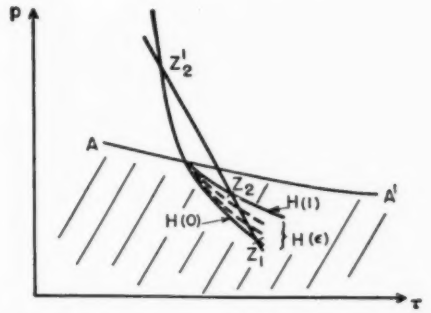


FIG. 3. Schematic drawing of one-parameter Hugoniot curves for steam.

type of result obtained. AA' represents the coexistence line, below which two-phase equilibrium prevails. The dotted curve is a dry-steam isentrope, continued metastably into the shaded coexistence region to the condition Z_1. The Hugoniot curve H, corresponding to Z_1 and to equilibrium conditions downstream, can be shown to be convex and to exhibit a sharp rise in the absolute value of its slope at its point of intersection with the coexistence line. The Rayleigh line through Z_1 intersects H in the two points Z_2 and Z_2'; Z_2 is the weak detonation solution and is the one that corresponds to the mixture of droplets and vapor observed experimentally. If the expansion continues downstream of Z_2, it does so along a mixed phase adiabatic. To our knowledge, no direct observation has been made that the flow downstream of Z_2 is supersonic in the case of steam; in the case of the irreversible condensation of a small amount of water vapor in air, or of air itself, this has been confirmed by observation of Mach lines.[2,14] One can, however, calculate sound speed at Z_2 with available thermodynamic data, obtain flow speed from the Hugoniot diagram (knowing the speed at Z_2), and establish that the flow is indeed supersonic.[15] Jouguet's rule[7] is thus verified. A Chapman-Jouguet process would occur in the special case in which the Rayleigh line was tangent to H at Z_2.

Figure 3 shows schematically a one-parameter family of Hugoniot curves representing idealized progressive stages of condensation advancing from the "dry" Hugoniot $H(\epsilon=0)$ towards the equilibrium Hugoniot curve $H(\epsilon=1)$. It is taken for granted that a single significant parameter ϵ may be found corresponding to a given mixture of growing molecular clusters (size larger than critical) and single molecules.[12] Here the family of Hugoniot curves is a pencil coalescing because of the single-phase condition above the coexistence line AA'.

This diagram suggests that the sequence of a weak

[14] D. Coles, Guggenheim, California Institute of Technology, private communication (1951).

[15] This can be verified, e.g., for the experiments reviewed by R. Ruedy, Can. J. Research **A22**, 77 (1944) [see references 14 and 23; also S. G. Reed, Jr., J. Chem. Phys. **20**, 539 (1952)]. There is question as to the value of sound speed in the two-phase mixture; whichever one chooses, however, in these cases the downstream Mach number remains supersonic.

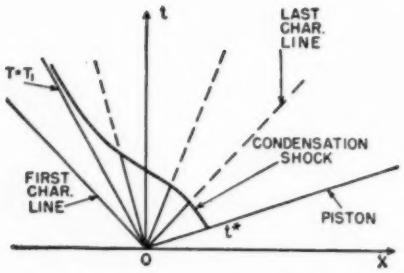


FIG. 4. Condensation shock in a centered simple wave expansion.

detonation proceeding from Z_1 to Z_2 and a normal shock transition leading from Z_2 to Z_2', will give the same end result as a "dry" normal shock leading directly from Z_1 to Z_2'. It can be verified easily that this "transitive" relation of the shock solution holds in general.

CONDENSATION SHOCKS IN ONE-DIMENSIONAL UNSTEADY FLOW

In two-dimensional steady nozzle flow, *normal* condensation shocks have been observed only within a limited domain of Mach numbers. There is, however, no reason why, in strictly one-dimensional flow, such shocks could not arise at any supersonic Mach number. In the following, some considerations are offered to support this view. It is desirable to ascertain the existence of normal condensation shocks on a broad basis since they are equivalent to chemical weak detonation processes, the occurrence of which is dubious. In order to obtain a variation of Mach number, we deal with unsteady flow as may be produced, e.g., by a piston receding with uniform velocity U from a semi-infinite air chamber. The air is first taken as dry; its initial state is denoted by the subscript 0. In an $x-t$-diagram the flow process is characterized by a pencil of striaght rarefaction lines [see Fig. (4)], along each of which the flow state is constant. This centered simple wave will be defined by the condition $u+\sigma=\sigma_0=\text{const}$ where u is the local particle velocity and σ is the Riemann function which, in the case considered here of an ideal gas, is known to be proportional to the local velocity of sound, $\sigma=2a/(\gamma-1)$. The equation of any rarefaction ("characteristic") line is then given as $x/t=u-a=k$, where $1/k$ is the slope of the line in the $x-t$-diagram.

Abandoning now the assumption of dry air, we stimulate the presence of a certain amount of water vapor, which may not yet be saturated at stagnation conditions and thus partakes in the expansion process. Suppose that at $T_1 \leqq T_0$ the vapor is saturated. In general, condensation will not occur at that temperature, since the change of state in the expanding gas column is not infinitely slow. A gas layer that became saturated with water vapor at time t_1 will continue to expand at constant entropy until, at some later time t, it attains a temperature $T<T_1$ at which condensation takes place. Let T_{sat} denote the saturation temperature corresponding to the static pressure prevailing in the gas layer at time t. Obviously, $T<T_{sat}<T_1$. The condensation delay $(t-t_1)$ will decrease if the difference between T_{sat} and increases, so that an expression of the form

$$t-t_1=F\left(\frac{1}{T_{sat}-T}\right) \tag{4}$$

will be expected to describe the general relationship of elapsed time and undercooling.[16] As $T\to T_{sat}$ this relationship may require refinement, as will be pointed out later.

The location taken by the gas layer when it reaches the temperature T evidently indicates the momentary shock location. The ordinate is known to be given by

$$\frac{t}{t_1}=\left(\frac{T_1}{T}\right)^{(\gamma+1)/2(\gamma-1)}, \tag{5}$$

whereas the abscissa follows from

$$\frac{x}{t}=u-a=\sigma_0\left[1-\frac{\gamma+1}{2}\left(\frac{T}{T_0}\right)^{\frac{1}{2}}\right], \tag{6}$$

since this is the equation of the characteristic line along which the temperature T prevails.

Every individual gas layer is characterized by the time, t_1, at which the air in it becomes saturated. On eliminating t_1 from Eqs. (4) and (5) we obtain

$$t=\frac{1}{1-\left(\frac{T}{T_1}\right)^{(\gamma+1)/2(\gamma-1)}}F\left(\frac{1}{T_{sat}-T}\right). \tag{7}$$

This is the ordinate of the shock path in terms of the parameter T which, in the $x-t$-diagram, determines the individual characteristic lines. The abscissa is given by Eq. (6).[17]

It is seen that the shock never can be found to the left of the characteristic line $T=T_1$ (since t would become negative if $T>T_1$); indeed, as $t\to\infty$, if $T\to T_1$, the shock path, coming from the low pressure side at right, approaches asymptotically the "saturation characteristic" $T=T_1$. Moreover, it can intersect but once with any characteristic line, since Eq. (4) is a unique function of T. As a consequence, the shock path, at any point P within the fan, will make a larger angle with the positive x-axis than does the rarefaction line through P. Hence, the shock velocity relative to

[16] A. Kantrowitz, J. Chem. Phys. **19**, 1097 (1951), derives the approximate formula

$$t-t_1=\frac{B}{(T_{sat}-T)^4};$$

also K. Oswatitisch, ZAMM **22**, 1 (1942).

[17] In the presence of the shock the dashed parts of the originally straight rarefaction lines (Fig. 4) will alter their course.

the tube wall, V, satisfies the condition $V<k=u-a$ at P, as is readily verified from the fact that the slopes of the intersecting lines are $1/V$ and $1/k$, respectively. The Mach number relative to the shock is $M>1$. If we now follow the shock curve as it approaches the asymptotic line, the difference between V and $k=u-a$ becomes increasingly smaller. It is thus seen that normal condensation shocks can occur theoretically at supersonic Mach numbers as low as we may choose.

On the other hand, the highest Mach number presumably will be found at the intersection of the shock with the last characteristic line on which any quantity will be given the subscript 2. The particle velocity here is equal to the piston velocity $u_2=U$. The shock velocity V can be obtained from the shock path within the zone of constant state extending between the piston path and the last characteristic line. In this zone, the right side of relation (4) is constant so that $t-t_1=t^*=\text{const}$. In the idealized expansion process considered here the gas layer immediately adjacent to the piston is saturated at time $t_1=0$; t^* then denotes the time at which condensation occurs in that layer.[18] The condensation shock makes its first appearance at $t=t^*$, $x=Ut^*$, which is a point on the piston path.

It is not hard to deduce that the path in the constant-state zone of a gas layer that was saturated at $t=t_1\neq 0$ is the straight line

$$Ut-x=a_2t_1\left(\frac{\sigma_1}{\sigma_2}\right)^{(\gamma+1)/(\gamma-1)}.$$

Since $t_1=t-t^*$, the shock path in that zone is given by

$$Ut-x=a_2t_2\left(\frac{\sigma_1}{\sigma_2}\right)^{(\gamma+1)/(\gamma-1)}.$$

This again is a straight line. Its reciprocal slope defines the shock velocity:

$$\frac{dx}{dt}=V=U-a_2\left(\frac{\sigma_1}{\sigma_2}\right)^{(\gamma+1)/(\gamma-1)}.$$

On introducing the simple-wave relation $\sigma_0=U+\sigma_2$, the Mach number relative to the shock in the constant-state zone (which includes the last characteristic line) can be written as

$$M_2=\frac{U-V}{a_2}=\left(\frac{\sigma_1}{\sigma_0-U}\right)^{(\gamma+1)/(\gamma-1)}$$

σ_1 and σ_0 are given quantities. It is seen that the Mach number M_2 becomes as large as may be desired if the piston velocity approaches the velocity

$$\sigma_0=\frac{2}{\gamma-1}a_0.$$

[18] Relations such as Eq. (4) cannot be expected to hold for very small times t^*. F. R. Gilmore, thesis, California Institute of Technology (1951), has estimated these induction times; for conditions of interest here they are of the order 10^{-8} sec.

The results of these considerations indicate that the apparent restriction of normal condensation shocks to a limited region of Mach numbers, as suggested by available experimental evidence in two-dimentional nozzle flow, is probably not founded in the nature of the irreversible condensation process.

Condensation shocks in unsteady one-dimensional flow have been obtained in shock tube experiments made at the University of Toronto.[19] Figure 5 gives an example showing the beginning expansion of initially saturated air. The situation here is evidently not that of a centered simple wave such as is depicted on the theoretical Fig. 4. Since the contact surface is not recognizable on Fig. 5, one cannot determine the thermodynamic state or the characteristics.[20] The condensation shock path, in such circumstances, need not necessarily seem to approach the first rarefaction line in the time available on Fig. 5.

However, experiments made by C. T. R. Wilson[21] suggest that this line may not even represent the true asymptotic of the shock path in the late stages of the expansion. Wilson found that before the onset of rapid condensation a minimum amount of undercooling is required in a slow expansion of saturated air. It appears that the condensation never will take place

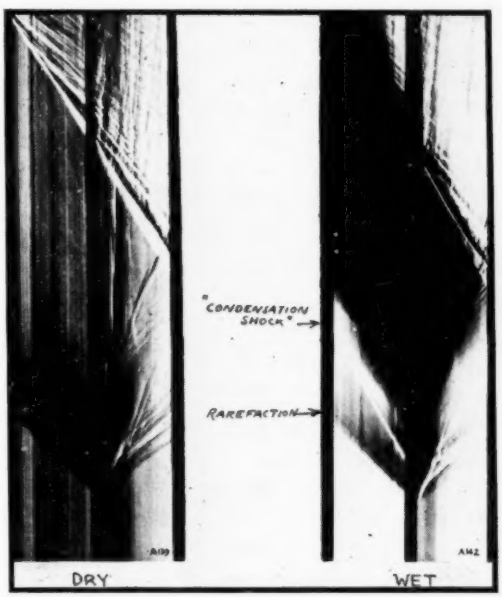


FIG. 5. Pictures showing flow processes in dry and wet air in a shock tube (courtesy of the University of Toronto, Institute of Aerophysics).

[19] The authors are indebted to Dr. G. N. Patterson and Dr. I. Glass of the Institute of Aerophysics, University of Toronto, for doing these experiments at our request, and for permission to publish one of the resulting wave-speed camera photographs.

[20] Possibly modification of these experiments is feasible in which the gas is allowed to push a piston, the path of which could be measured [Michels, Slawsky, and Jacobs, Physica **XX**, 223 (1954)]. With this the condensation history from time zero could be observed.

[21] C. T. R. Wilson, Proc. Roy. Soc. (London) **A61**, 240; Phil. Mag. **189**, 281 (1897).

at the equilibrium saturation temperature T_1 which therefore should be replaced by the "Wilson temperature" in Eqs. (5) and (7). Correspondingly, the asymptotic line approached by the shock path in Fig. 4 should be replaced by the "Wilson characteristic."

One can show that, for a centered simple wave expansion, the condensation path should appear quite straight during a sufficiently short time interval. Figure 5 suggests that the path may be nearly straight also in a more general expansion process when the time is measured in fractions of milliseconds.

It is interesting to remark that Lee[22] has observed three-dimensional unsteady condensation shocks due to jet formation when a diaphragm is burst in a shock tube. These would correspond to the weak detonation solution indicated by G. I. Taylor.[9]

DISCUSSION OF SOME PERTINENT DETAILS OF DETONATION PROCESSES

A more detailed description of the weak condensation detonation processes discussed above may help to clarify their relation to more familiar types of chemical detonation processes.

If one tries to describe the processes occurring in the detonation front itself, the theory of steady detonation waves (see, e.g., von Neumann,[8] Kirkwood and Wood[23]) gives some support to the notion that ordinary chemical detonations start from a situation in the neighborhood of a normal shock solution of the conservation equation for the unburned gas and proceed along a sequence of (local) steady states to the Chapman-Jouguet point.

In the case of the weak condensation detonations, we have seen that the analogous sequence of states starts in the neighborhood of the "weak" disturbance solution and remains there.[24]

For both sorts of detonation the equilibrium of molecular species or phases is reached in a kinetic balance of forward and reverse reactions. In the usual chemical process, both forward and reverse reaction rates have positive partial derivatives with respect to temperature. In the case of condensation processes the corresponding derivative for the forward reaction rate (the condensation) is negative while that of the reverse reaction rate (evaporation) is positive. The process may cut itself off (i.e., equilibrium may be established) even if some of the vapor is not yet condensed.

In order to obtain the weak condensation detonation, an unstable condition is brought on by means of a fast adiabatic expansion. Two basic mechanisms are at work in such a condensation process. The first stage, as is well known, depends on the buildup of "critical" nuclei; once these are present in sufficient quantity the rapid stage of condensation commences. The first stage is "hidden" insofar as it is not made manifest by any appreciable alteration of thermodynamic properties. The evaporation (reverse process) depends on the presence of droplets and on the heat release due to their fast initial growth.

This description prepares the way for a comparison with the weak detonation solution found by Friedrichs.[10] Friedrichs has treated the one-dimensional model detonation process as governed by a differential equation involving macroscopic viscosity, heat conduction, and a reaction rate. He has shown the possibility of weak detonation solutions of these equations and has shown that these will have an eigenvalue of locally steady propagation velocity for given boundary conditions ahead of the wave and for a given form of the reaction rate. In this, these solutions have formal similarity to flame solutions. Assuming a reaction rate increasing with temperature, the weak detonation solutions are characterized as being associated with a reaction rate which he terms "extremely fast." Much of his discussion is independent of this assumption. What is essential to the matter is the behavior of the solutions of his model equations in the $(\epsilon_1, \vartheta_1, \tau)$-space, where ϵ is the fraction of burnt gas, ϑ = a reduced temperature, and τ = specific volume. To have a weak detonation solution, Friedrichs finds, in accordance with physical expectation, that the transition between the planes $\epsilon=0$ to $\epsilon=1$ must be effected without departing far from the plane $\vartheta=\vartheta_1$, where ϑ_1 is the reduced temperature ahead of the reaction front. This is the reason for the requirement that the reaction rate be fast at a relatively low temperature.

It is known from experience that the weak condensation detonations satisfy the condition of transition from $\epsilon=0$ to $\epsilon=1$ without marked increase in temperature. The reaction rate has two fast stages: rapid growth on nuclei, and the reverse reaction of evaporation moderating the growth of droplets. A second type of weak detonation solution found by Friedrichs involves a shock transition in the neighborhood of the plane $\epsilon=1$, i.e., in the burnt gas. This would not seem to be possible in the condensation case because of the stabilizing conditions of the reaction rate.

CONCLUDING REMARKS

The rapid condensation processes described above are examples of weak detonation transitions. They show that such transitions actually occur. That they are not initiated by a shock, is the decisive difference from the usual chemical detonations. Friedrichs has shown the possible existence of weak detonations under the condition of not being initiated by a shock. However, in chemical processes weak detonation transitions do not seem to have been observed.

[22] J. Lee, dissertation, University of Toronto (1950).

[23] J. G. Kirkwood and W. W. Wood, J. Chem. Phys. **22**, 1915 (1954).

[24] The fact that the disturbance is small in the condensation case indicates that one might compute eigenvalues of the propagation velocity by an appropriate modification of a procedure used for flames [D. Layzer, J. Chem. Phys. **22**, 222 (1954)]; inversely, the pattern of solution fields for different assumptions as to the reaction rate indicates more closely what conditions might have to be met to have a chemical weak detonation.

Preparation of Thin Magnetic Films and Their Properties*

M. S. Blois, Jr.†
U. S. Naval Ordnance Laboratory, Corona, California
(Received February 7, 1955)

Techniques for the preparation by vacuum evaporation of single thin layers and laminated structures of ferromagnetic alloys (principally in the iron-nickel system) are described. The magnetic properties are related to those parameters of the process which may be chosen to yield materials having desired characteristics. A method for producing films with extreme hysteresis loop rectangularity is described. Questions of suitable substrate materials, evaporable dielectrics for use in insulating multilayer deposits, and control of alloy composition, are considered in detail.

INTRODUCTION

THE wide, present-day interest in solid-state electronic components has resulted in a number of new developments in magnetic materials. Many of these developments are concerned with nonmetallic substances such as the ferromagnetic ferrites; others are improvements in the fabrication of the older and better known magnetic alloys. The present research was undertaken to learn if the improved magnetic alloys, prepared in extremely thin sheet form, could be used to advantage as magnetic-core materials for high-frequency applications. The first problem encountered in the use of magnetic materials at higher frequencies is that of increased eddy-current losses. The ordinary solution to this problem is to laminate the core material. This approach has apparently reached its limit in the application of thin-rolled ribbon alloys, which are now available in thicknesses down to approximately 10^{-4} inch, where eddy currents are negligible at frequencies below 10 mc. We have proposed to reduce the thickness of alloy sheets even further through the use of a vacuum evaporation process, depositing both the magnetic alloy and an insulating material *in situ* to build up alternating layers of magnetic and dielectric films.

It was expected that other worthwhile properties might ensue. For example, in a digital computer using a static magnetic-core memory based on Forrester's arrangement,[1] it is essential that the hysteresis loop of the magnetic material used be sufficiently rectangular. It has been shown[2] that if a ferromagnetic layer is sufficiently thin (perhaps 3000 A or less), it is energetically unfavorable for the layer to have domain boundaries in the absence of an applied magnetic field. Such a thin layer will then consist of a single magnetic domain, and its hysteresis loop may be quite rectangular.

It was also hoped that the vacuum-evaporation method might prove useful as a production technique in specific applications, as in the manufacture of new computers that might require enormous numbers of storage elements of the order of one-hundred thousand or more.

VACUUM EVAPORATION OF MAGNETIC FILMS

The preparation of evaporated metallic films with thicknesses in the range from 100 A to 10 000 A has been widely studied and, more particularly, a smaller amount of work on magnetic films has been reported.[3] Certain problems arose in our work, however, which are not normally encountered in thin-film studies, and these are worthy of mention.

The choice of substrate was considered to be of primary importance, not only to insure that the surface would be microscopically smooth but also to provide compatible values for the thermal expansion coefficients of the film and the substrate. The choice of dielectric and the technique to be used in its evaporation required evaluation. The question of the extent to which a molten alloy would fractionate during evaporation had not been studied for the alloys in question, although it had been shown in the case of brass that if evaporation was carried out quickly enough the evaporated alloy film would closely resemble the starting material, even in crystallographic detail.[4] Since the magnetic behavior of the ferromagnetic alloys may be strongly dependent on thermal history, a correlation between the principal magnetic properties and several different annealing procedures was also sought.

The vacuum system and heating arrangements used were conventional,[5] the vacuum ranging from 5×10^{-5} mm Hg to 10^{-3} mm Hg, and sometimes momentarily higher, during the course of evaporation. Because of the tendency of molten iron, nickel, and cobalt to alloy with tantalum, molybdenum, and other refractory metals commonly used for crucibles, it was necessary to turn to the use of pressed alumina crucibles and high-frequency induction heating. With a commercial induction heating unit having a nominal power output

* This research has been supported by the Office of Naval Research.
† Now at the W. W. Hansen Laboratories of Physics, Stanford University, Stanford, California.
[1] J. W. Forrester, J. Appl. Phys. **22**, 44 (1951).
[2] C. Kittel, Revs. Modern Phys. **21**, 541 (1949).
[3] The work of Crittenden and associates at Case Institute of Technology appears to be the most recent. See Revs. Modern Phys. **25**, 310 (1953).
[4] L. Harris and B. M. Siegel, J. Appl. Phys. **19**, 739 (1948).
[5] See, for example, J. Strong, *Procedures in Experimental Physics* (Prentice-Hall, Inc., New York, 1945).

of 2 kw, a sample of iron or iron alloy of 10–15 grams could be melted quite conveniently.

A spontaneous peeling or tearing of the film is a common difficulty experienced in attempting to form relatively thick films, say, of 5000 to 10 000 A. This indicates that the film exists in a state of high stress which might be attributed to differential thermal expansion between substrate and film or to stacking faults in the lattice created during the rapid condensation of the metal atoms into a solid. It was found that this stress condition could be reduced by raising the temperature of the substrate during deposition of the film. Most of the films were consequently made with the substrate mounted in a heater which maintained a predetermined, elevated temperature.

SUBSTRATES

The first requirement was that the substrate be smooth on a scale comparable with the film thickness. That is, the substrate surface need not be flat in the optical sense, but while it may be undulatory in profile there should not be any small-scale scratches, pits, or bumps. This requirement is easily met in the case of glass-like materials which may be fire polished, but there was at the beginning the additional requirement that the substrate should withstand fairly high annealing temperatures, up to perhaps 1000°C.

A priori it seemed that the coefficient of thermal expansion of the substrate should match that of the alloy, since the condensed film is formed at an elevated temperature, and stresses set up by the cooling of the structure to room temperature might result in undesirable magnetic properties.

Initial efforts were therefore directed toward finding glazed ceramic materials with very high expansion coefficients (of the order of 12×10^{-6} per degree C). Unglazed ceramic bodies, no matter how carefully ground or polished, are unacceptable because of their relative roughness on a microscopic scale.

During this search it was found that ordinary microscope slides would serve fairly well at moderate temperatures. Their coefficient of expansion, which is probably in the range $8–10\times10^{-6}$ per degree C, compares well with that of the 50% Ni–50% Fe alloy, which is about 9.7×10^{-6} at comparable temperatures.

The nature of the forces existing between the metal atoms at the interface and the substrate is relevant here, and, as in the case of other thin-film studies, it is found to depend strongly upon the substrate and its state of cleanliness. The principal forces involved are usually considered to be of the van der Waals type, but it seems likely that chemical bonds may be formed in some instances, giving a much greater stability; for example, thick iron-alloy films which have been deposited on heated glass and then subjected to violent thermal shock have been observed to pull off, taking a thick layer of glass with them. This may be an example of chemical binding to the interstitial or bound oxygen in the glass, or alternatively, the result of diffusion of metal atoms into the glass providing a sort of atomic dovetailing.

DIELECTRICS

Among the requirements to be met by a dielectric for use in magnetic-film lamination are evaporability, chemical inertness with respect to the metal layers at high temperatures, and stability at these temperatures, in addition to the necessary insulating properties. The first dielectric tried was silicon monoxide, which has come into common use as a substrate material for electron microscopy because of the ease with which it may be made into thin, evaporated films. Although it formed satisfactory films, during prolonged annealing it reacted with the thin metal films and spoiled them. Similar behavior has been reported by other workers, and it appears that either residual silicon or oxygen in the material may be responsible.

High purity magnesium fluoride of the grade used for the coating of optical elements was employed with some success. This material, as well as the silicon monoxide, was evaporated satisfactorily from a tantalum boat by resistance heating. It appeared to be chemically inert at temperatures up to 1200°C, and while trouble was occasionally experienced with the multilayered deposits tearing and wrinkling during annealing, the dielectric did not appear to be at fault. Whenever peeling of the laminated films was observed, it always occurred between the bottom metal layer and the substrate rather than between laminations, and thus it appears that the metal-to-dielectric binding is stronger than the metal-to-substrate binding.

COMPOSITION OF EVAPORATED ALLOYS

Harris and Siegel, in the paper previously referred to, reported on the evaporation of α and β brass. It was found that if the rate of evaporation was sufficiently high, the resulting films would be both chemically and crystallographically identical with the starting material. This approach is the reverse of that employed in fractional distillation; i.e., if the alloy mixture is heated to a very high temperature and evaporated quickly, the equilibrium indicated by the phase diagram is never reached.

Using the vapor-pressure data for metals given by Dushman,[6] and applying the law of Raoult for the vapor pressure of solutions, one may determine the relative rates of evaporation for each constituent of a given alloy at any given temperature. Thus, for a typical situation of interest, with a 50-50 iron-nickel alloy (Hipernik, Deltamax, etc.) maintained at 1600°C, the vapor pressure, and hence the evaporation rate, of the iron is about three times that of the nickel, so that the resulting film would be expected to have the

[6] S. Dushman, *Scientific Foundations of Vacuum Technique* (John Wiley and Sons, Inc., New York, 1949).

composition 75% Fe, 25% Ni. In the case of 4–79 Mo-Permalloy at the same temperature, the nickel evaporates about 50% faster than the iron, but the molybdenum scarcely evaporates at all, the vapor pressures being approximately 27μ, 18μ, and $1.2\times10^{-3}\mu$, respectively.

These predictions are based upon the validity of applying Raoult's law to the existing situation. Preliminary experimental results indicate that the difference in composition between starting material and evaporated film is considerably less than would be expected from the theory above, at least in the case of the iron-nickel alloys studied thus far. For the nickel determination, the film is dissolved in nitric acid, oxidized to the higher valence state with bromine, and neutralized with ammonium hydroxide. This also removes the excess bromine color. Addition of dimethylglyoxime produces the red nickel complex, the optical density of which is measured in a Beckman spectrophotometer and compared with standard solutions. Dilutions are kept high enough to stay within the linear (Beer's law) range. The amount of iron is then determined on a separate aliquot by a thiocyanate reaction in much the same way.

The results for one set of 9 films evaporated sequentially from a single crucible charge are shown in Fig. 1. The composition of the starting material prior to the run was 16.6% Fe–83.4% Ni, and at the conclusion of the run was 16.4% Fe–83.6% Ni with approximately $\frac{4}{5}$ of the crucible charge remaining. For this composition and a temperature of 1600°C, according to the method previously described, one calculates vapor pressures for the iron and nickel of 17.2μ and 28.1μ, respectively, which would yield a film of approximate composition 38% Fe–62% Ni. The films actually had a mean composition of 18.5% Fe–81.5% Ni, which showed the expected tendency of enrichment in iron, but by a much smaller amount. The preparation of other alloys in this system is now under investigation and findings will be reported later.

EFFECT OF ANNEALING

It was found that the coercivity of the films was strongly affected by the nature of the substrate and that, with most substrates, the coercivity depended upon the substrate temperature at the time of formation as well. This was assumed to be a matter of residual strain or lattice disordering, and it was hoped that further heat treatment would shed additional light on this. It was found that when 45% Permalloy films were formed on several different glazed ceramic substrates at room temperature the coercivity was initially of the order of 20 oersteds. Annealing in dry hydrogen at 800–1000°C for 6 hours would reduce this in a typical instance to approximately 10 oersteds. In contrast, films of 80-Permalloy exhibited initially relatively low coercivities which increased substantially upon annealing.

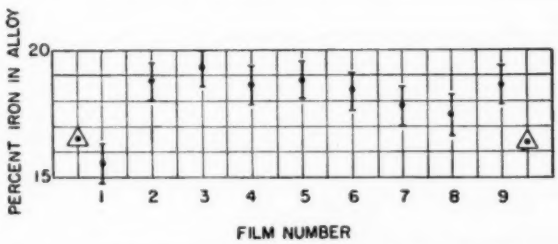


FIG. 1. Chemical composition of a series of films approximately 10 000 A thick, evaporated from an alloy with original composition 16.6% Fe–83.4% Ni. ● Composition of evaporated film; △ Composition of crucible charge.

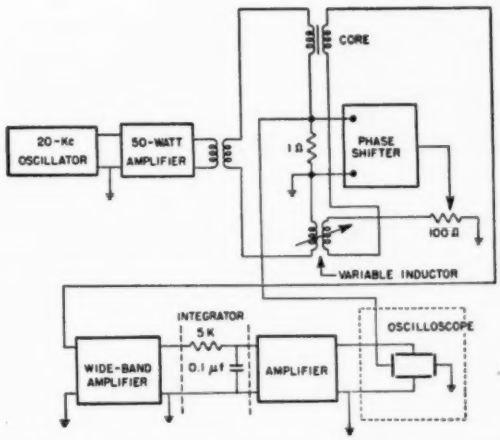


FIG. 2. Schematic of the hysteresis loop display equipment. The film under test is used as the core of the mutual inductance shown.

One series of films of 50% Fe–50% Ni showed upon annealing an average reduction in coercivity of 48%. Experimental data reported by Bozorth[7] on the anistropy constant (to which the coercivity is proportional) shows a value of 1.0×10^4erg/cm³ for this alloy in the disordered state in which the freshly formed films may be considered to be, and a value of about 0.6×10^4 erg/cm³ for the ordered state. Thus, if the coercivity depended only on the anistropy, we would expect it to be reduced during annealing by 35%. The difference between these values, which has considerable experimental uncertainty, suggests that with the 50% alloy a large part of the reduction in coercivity upon annealing may be due to a decrease in anistropy, any remainder owing to a reduction or elimination of strain.

With 80% Permalloy, where the ordered state has a greater anistropy (increasing negatively), the annealing process was found to increase the coercivity, typically from about 0.5 oersted to 3.3 oersteds. This increase by a factor of 6.6 is compared with an increase of about 4.6 predicted on the basis of increased anisotropy alone. This difference is probably not significant for the data at hand.

It has been mentioned earlier that spontaneous peeling of thicker films could be avoided by maintaining

[7] R. M. Bozorth, Revs. Modern Phys. **25**, 42 (1953).

FIG. 3. The hysteresis loop of a 10 000 A thick 80-Permalloy film (H_c=0.48 oe). It is typical of films of this composition to have a small coercivity and little tendency toward rectangularity in the absence of further treatment.

the substrate at an elevated temperature during deposition. This effect is believed to result from a reduction of lattice misfits during the growth of the microcrystals because of the enhanced mobility of the semifree surface atoms at higher temperatures. The control of evaporation rate and film-forming temperature should accomplish much the same result as annealing; and in dealing with the structure of evaporated films, one presumably must consider the process of film formation as a part of the annealing history.

We have found a temperature of about 300°C to be a minimum satisfactory substrate temperature for a typical ceramic substrate material and a Ni–Fe alloy. If the mechanism proposed above is correct, then this temperature should be independent of substrate material once the initial layers are formed and depend mainly on the activation energy for migration of the iron and nickel atoms over the freshly formed alloy surface. By maintaining the substrate at or above this minimum temperature, we find that films can be satisfactorily formed. Furthermore, it has been found that by suitably controlling the substrate temperature we can avoid the necessity of subsequent annealing.

Using the evaporation rate of approximately 10 000 A/min, an 80-Permalloy film formed on a substrate at 350°C will have a coercivity of the order of 1–2 oersteds. The film may then be annealed and slowly cooled to produce an ordered structure which will show a coercivity of about 10 oersteds. On the other hand, if we employ a higher substrate temperature, say 550°C, and the same evaporation rate, we find the fresh film has a coercivity of about 10 oersteds. The increased mobility during film formation has presumably resulted in a fairly high degree of order. We have therefore concluded that the order-disorder state of the material can be influenced during the evaporation process by properly adjusting the rate of evaporation and the substrate temperature.

EXPERIMENTAL RESULTS

The magnetic properties of the resulting evaporated films were measured at 60, 20 000, and 10^6 cps. A hysteresis loop and the unintegrated voltage output were photographed for each specimen both before and after any annealing—and with some specimens after various intervals of time to see if there were any ageing effects. The general method used is similar to that described by Crittenden[8] and involves two mutual inductance coils connected with their outputs in opposition so that coupling through the air is canceled out. The magnetic film is inserted in one inductance, and the resulting voltage unbalance is then due essentially to the film alone. This voltage is integrated, adjusted in phase, and displayed on an oscilloscope as B. The voltage developed across a resistance in series with the magnetizing current is connected to the horizontal plates of the oscilloscope to give H. Figure 2 shows the general arrangement used for the 60-cycle and 20-kc measurements. For the 60-cycle measurements the coils are approximately 2 inches in diameter by 6 inches in length. Films were introduced into the field near the axis of the solenoid by use of a suitable sample holder so that the plane of the film was parallel to the axis. It is interesting to note that for plane films magnetized along an axis in the plane of the film, demagnetizing effects are negligible owing to the extreme thinness of the films.[9] Measurements of the hysteresis loops of cylindrical and annular films in the circular direction were also made at both 20 kc and 1 mc by winding torroidal coils on the films. Similar windings were put on an uncoated ring to compensate for the air coupling, and a small adjustable mutual inductance was used as a fine adjustment to achieve cancellation of the directly coupled signal. Figures 3–6 illustrate typical properties for the materials and parameters shown.

FIG. 4. This film of 45-Permalloy (10 000 A thick) on a steatite substrate, illustrates how the coercivity may be reduced by annealing, with a loss in rectangularity. Loop at left is unannealed (H_c=9.2 oe) and the annealed loop is shown at the right (H_c=3.6 oe).

(A) (B) (C)

FIG. 5. Curve (A) is the hysteresis loop of a 45-Permalloy film (10 000 A thick) immediately after preparation (H_c=13.8 oe). Curve (B) is after a 7-hour anneal in hydrogen at 450°C followed by an air quench (H_c=4.1 oe). Curve (C) shows the loop after a further 6-hour anneal at the same temperature and air quenching. The coercivity has now risen to 7.8 oe. See also Fig. 7.

SPEED OF REVERSAL OF MAGNETIZATION

The magnetic properties of greatest interest for computer application are the hysteresis loops at 60 and 20 000 cycles and the reversal time. If the reversal of polarization takes place by domain-wall motion, the velocity of which is nearly proportional to $H-H_c$ where H is applied field and H_c is coercive force, then the reversal time should be inversely proportional to the

[8] E. C. Crittenden, Jr., *et al.*, Rev. Sci. Instr. **22**, 872 (1951).

[9] See, for example, R. M. Bozorth, *Ferromagnetism* (D. Van Nostrand Company, Inc., New York, 1951).

same quantity. The electron spin relaxation time sets an upper limit to the wall velocity—that at which the wall with its finite width crosses a fixed lattice point in one relaxation period. Reversal of polarization by a domain rotation process in which all spins rotate more or less uniformly may be faster, however, since the entire specimen may be expected to reverse in as little as one relaxation period. It was hoped that a study of these reversal times as a function of field for films of differing thickness would give a clue to the reversal process.

Figure 7 shows the reciprocal reversal time as a function of field for several magnetic materials including an evaporated, laminated 80-Permalloy film of seven layers of approximately 10 000 A each. The curves of the rolled alloy and ferrite material are taken from the data reported by the Digital Computer Laboratory of Massachusetts Institute of Technology and the Armour Research Foundation.

The very similar performance of the evaporated material and the thinnest rolled alloy supports the view that at these thicknesses eddy currents play a minor

FIG. 6. The loop at left is of a fresh eight-layer deposit of 80-Permalloy (MgF_2 insulation) having a total metal thickness of approximately 79 000 A as determined by weighing. The loop at the right is the same deposit taken 3 days later without any intervening treatment. It appears that diffusive or other spontaneous lattice rearrangements may have relieved some of the internal strains and resulted in greater rectangularity.

role, and the switching time is limited largely by fundamental magnetic processes.

The reversal time was defined for the evaporated films as the width of the output signal at half amplitude and was measured by a pulse technique such that the value of the driving field was constant during the reversal process.

It should be added that in the matrix-memory application using the coincident-current selection method of Forrester the maximum driving field is limited to twice the coercive field. Referring to Fig. 7, the 4–79 Mo-Permalloy with an H_c of 0.1 oersted will then require about 5 μsec to reverse under a driving field of 0.2 oersted. The evaporated material with a higher H_c of 0.7 oersted will reverse in about 0.5 μsec with a driving field of 1.4 oersteds.

HYSTERESIS LOOP SQUARENESS

Theoretically a sufficiently thin evaporated magnetic film is expected to be single domain and might therefore be hoped to have a rectangular loop. The single-domain character of very thin films has been experimentally

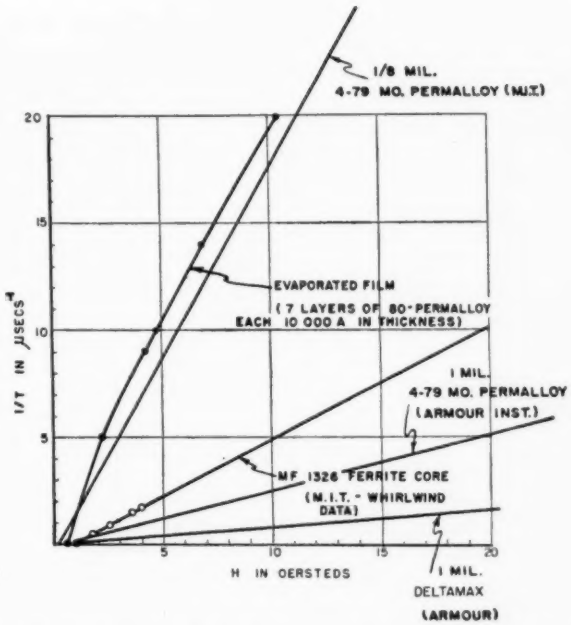


FIG. 7. Reciprocal polarization-reversal time as a function of driving field for typical high speed materials.

verified by Drigo and Pizzo[10] and Crittenden.[11] Under some circumstances, however, one might prefer to be able to deposit thicker films and still obtain rectangular loop behavior.

It was suspected that an orienting effect might be achieved by allowing the metallic vapor to condense on a substrate placed in a magnetic field of proper strength. The hysteresis loop of a 1850-A film with composition 80% Ni−20% Fe is shown in Fig. 8. This film has the configuration of an annulus with an o.d. of 1.5 cm and an i.d. of 0.4 cm and was formed with a magnetic field of approximately 350 oersteds in the plane of the film. Substrate temperature was approximately 325°C. The left curve of Fig. 8 is the hysteresis loop measured with the driving field parallel to the preferential magnetic axis, the middle is measured perpendicular to this axis, and the right curve is taken at an angle of 45 degrees. The coercive force for the rectangular loop is about 1.6 oersteds. As noted in the foregoing, because of the one very thin dimension, the demagnetization is negligible and the magnetic properties are essentially independent of the complicated configuration used; i.e., the hole makes no observable difference.

FIG. 8. Hysteresis loop of an 80-Permalloy film, 1850 A thick deposited in an orienting magnetic field. From left to right, loops taken with the magnetizing direction parallel, perpendicular and at 45° to the axis of orientation.

[10] A. Drigo and M. Pizzo, Nuovo cimento **5**, 1 (1948).
[11] E. C. Crittenden, Jr., Revs. Modern Phys. **25**, 310 (1953).

FIG. 9. An oriented 6000 A thick film of 80-Permalloy deposited on Mylar. The hysteresis loops are due to parallel, 45°, respectively and perpendicular magnetization with respect to the axis of orientation.

Smaller fields have been found to induce orientation, and it appears that even the earth's field may introduce a detectable orientation. The degree of orientation appears to fall off however as the magnetic field strength is reduced, and to insure fairly complete orientation we have so far been using fields in the range 200–1000 oersteds. One might expect that the higher the substrate temperature, the higher the required field.

In comparison with other magnetically oriented materials, the evaporated films deposited in a magnetic field have the possible utilitarian advantage that they are extremely durable. They sometimes adhere so tightly to the ceramic substrate that they cannot be scraped off and must be filed if, for example, it is desired to subdivide a given region. The loop rectangularity that is achieved during deposition appears to be a property which is permanent under quite wide ranges of temperature and mechanical shock.

In addition to the deposits made on rigid ceramic substrates, the prospects of forming an improved magnetic material useful for magnetic tape or drum recording were examined by forming magnetically oriented deposits on a plastic substrate. Figure 9 shows the hysteresis loops of a 6000 A 80-Permalloy layer deposited on a Mylar substrate in the presence of a 400 oersted field. The resulting coercivity is about 7 oersteds.

Because of a preoccupation with the possibilities of employing these materials in the specific case of matrix-type memory systems, a major emphasis has been placed on achieving low coercive forces. Since we have been largely concerned with the different Permalloy compositions, we turned to a composition in the neighborhood of 75–80% nickel in the hope of simultaneously reducing both the anisotropy and the effects of residual strain acting through magnetostriction. The films with lowest coercivity have had an H_c of about 0.1 oersted, and the mean H_c for the 80-Permalloy films has been in the neighborhood of 1.0 oersted.

In view of the possible application of these methods to magnetic tape and drum recording and also as a matter of general interest, several attempts were made to produce high-coercivity films. These are to be distinguished from the many accidental high-coercivity materials resulting from highly strained films, which tend to be mechanically unstable. It was found that a cobalt film deposited in an orienting magnetic filed and formed with its single axis of easy magnetization in the direction of the applied field had a coercivity of approximately 40 oersteds.

SUMMARY

It has been found possible to prepare by a vacuum evaporation process thin ferromagnetic films and laminated structures thereof with individual thicknesses ranging from 1000 A to 100 000 A and with coercivities of 0.1 to 40.0 oersteds. A magnetic orientation method has been employed to produce rectangular loop materials for use in applications where this property is desirable. Microchemical methods of controlling film composition have made possible the production of iron-nickel alloy films having the desired composition. Magnetic properties have been found to be uniform to within a few percent for a single batch of cores, but further work appears necessary to improve the batch-to-batch reproducibility.

The present state of knowledge, though by no means complete, appears for certain materials and configurations to offer a reasonable point of departure for development work aimed at the practical utilization of materials of this type.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the many contributions of G. Miles and R. Fall in the preparation of these materials, and of R. Conger and J. Hensey in the measurement of their properties.

On the Reflection of Shock Waves from an Open End of a Duct*†

GEORGE RUDINGER
Cornell Aeronautical Laboratory, Inc., Buffalo, New York
(Received August 21, 1954)

The readjustment of the mean exit pressure to its steady-flow equilibrium level, following the arrival of a shock wave at an open end of a duct, has been investigated. For a cylindrical duct in which the gas is initially at rest, a solution based on acoustic theory was obtained in the form of a Fourier integral. The nature of this solution was established as a series of damped harmonic oscillations superposed on an exponential pressure decay. The actual pressure decay curve was obtained by numerical integrations. The same problem was also attacked experimentally using a shock tube of 3.23 inches i.d. From pressure records taken some distance from the open end, the "effective" exit pressure was derived.

It was found that the exit pressure deviates appreciably from its steady-flow equilibrium value during a time in which a sound wave could travel about three duct diameters. Satisfactory agreement between theory and experiment was observed for shock pressure ratios up to about 1.9 (in air) although, near this limit, significant deviations owing to wall friction were observed toward the end of the pressure decay region. For still stronger shocks, the influence of the decay time on the phenomena inside the duct becomes unimportant while friction effects dominate. The possible significance of the lag in the establishment of steady-flow boundary conditions in practical applications is discussed, in particular, for cases where the acceleration of the gas by the reflected expansion wave may be of importance.

I. INTRODUCTION

PROBLEMS of quasi-one-dimensional nonsteady flows in ducts are generally attacked by means of the method of characteristics, e.g., references 1 and 2. This method enables one to determine graphically the gas motion and the propagation of pressure waves in the time-distance plane from given initial and boundary conditions. When the end of the duct is closed, the boundary conditions require that the flow velocity there must be zero at all times. For an open end, however, the flow field does not terminate at the "boundary": Pressure waves pass through the duct exit and create an external, nonsteady flow field. The problem of analyzing the latter can be avoided if one makes reasonable assumptions about the flow conditions at the open end. Wave reflection can then be determined without consideration of the external flow phenomena which are implied in the assumptions made.

It has customarily been assumed that the boundary conditions in nonsteady flow are the same as in steady flow. Under this assumption, one neglects the time necessary for the steady-flow boundary conditions to be established after the disturbance caused by an arriving pressure wave. This time, which depends on the dimensions of the cross section of the duct, is generally small compared to the time in which the flow conditions change significantly in the absence of shock waves. Use of the steady-flow boundary conditions has, therefore, led to good results.[3] It is usual practice to apply the steady-flow boundary conditions also in the case of shock reflection from an open end, in spite of the fact that a shock changes the flow conditions practically instantaneously. This investigation was undertaken to determine the errors that may be introduced in a wave diagram as result of using steady-flow boundary conditions. Experimental evidence which pointed to the need for investigation of this problem was also obtained by Glass,[4] and Calhoun and Kogan.[5]

The investigation was carried out experimentally by means of shock tube techniques and theoretically on the basis of acoustic theory. In the latter approach, the reflected wave was determined from its frequency spectrum which, in turn, was derived from that of the incident wave with the aid of the specific acoustic impedance of an open-ended circular duct.[6] One might have expected that only pressure changes produced by extremely weak shock waves would be properly accounted for by acoustic theory. It was found, however, that the theoretical results are satisfactory at least

* This research was conducted under the auspices of Project SQUID, jointly sponsored by the Office of Naval Research, Department of the Navy, the Office of Scientific Research, Department of the Air Force, and the Office of Ordnance Research, Department of the Army.

† An abridged version of this paper was presented at the Second Canadian Symposium on Aerodynamics held at Toronto, Ontario, February 25–26, 1954.

[1] R. Courant and K. O. Friedrichs, *Supersonic Flow and Shock Waves* (Interscience Publishers, Inc., New York, 1948).

[2] G. Rudinger, *Wave Diagrams for Nonsteady Flow in Ducts* (D. Van Nostrand Company, Inc., New York, 1955).

[3] E. Jenny, Brown Boveri Rev. **37**, 447–461 (1950).

[4] I. I. Glass, "The design of a wave interaction tube," University of Toronto, Institute of Aerophysics, UTIA Rept. No. 6, May 1950.

[5] J. D. Calhoun and A. Kogan, "A preliminary investigation of the flow and reflection conditions associated with a uniform normal shock wave emerging from a two-dimensional sharp-edged channel," Princeton University, Aeronautical Eng. Lab. Rept. No. 209, June 1, 1952.

[6] Analytical solutions to problems involving pressure waves of finite amplitude have been published by a number of investigators. W. Chester, Trans. Roy. Soc. (London) **A242**, 527–556 (1950), studied the two-dimensional diffraction pattern produced by a sound pulse that reaches the end of a channel consisting of two semi-infinite parallel planes; R. E. Meyer, Quart. J. Mech. Appl. Math. **5**, (3), 257–291 (1952), obtained a solution for the flow in a duct if a disturbance at the end is prescribed and if the deviations of the flow variables from an originally steady flow are small; various methods of linearization were also used by K. O. Friedrichs, Comm. Appl. Math. **1**, 211–245 (1948), G. B. Whitham, Proc. Roy. Soc. (London) **A203**, 571–581 (1950), and others (see also references 7 and 8).

up to shock pressure ratios approaching 2.0. Similar conclusions were reached by Keller[7] and Hudson[8] who studied shock diffraction around obstacles by means of acoustic theory. Hudson estimated that shock pressure ratios of 1.6 represented about the limit of usefulness of the acoustic theory in his case.

II. FORMULATION OF THE PROBLEM

When a shock wave reaches an open end of a duct, the subsequent flow phenomena in the vicinity of the exit are highly complex as indicated by schlieren photographs,[9] streak camera records,[4] and interferograms.[5] All of these show the formation of a vortex ring which briefly remains stationary at the exit and is then swept downstream. There is also evidence[4] that waves which originate in the external region may re-enter the duct. All these phenomena are three-dimensional while, for the purpose of wave diagram construction, one requires "effective," one-dimensional boundary conditions. Instead of trying to extract these from detailed observations of the three-dimensional phenomena at the exit, an indirect approach was used: Measurements were taken at some distance from the exit, where the flow again may be considered as one-dimensional, making it possible to infer from these the effective exit conditions with the aid of wave diagram procedures.

For a derivation of wave diagram procedures and properties of characteristics, reference must be made to the literature on the method of characteristics, e.g., references 1 and 2. However, it may be recalled here that in the case of isentropic flow of an ideal gas in a duct of constant cross section, the characteristics are two families of curves whose slope in the x,t-plane (the wave diagram) is given by $dx/dt=u\pm a$ where u is the flow velocity and a the speed of sound. Along these lines, the following relations apply:

$$P=2a/(\gamma-1)+u=\text{const} \quad \text{for} \quad dx/dt=u+a \quad (1)$$

and

$$Q=2a/(\gamma-1)-u=\text{const} \quad \text{for} \quad dx/dt=u-a. \quad (2)$$

A wave diagram showing the reflection of a shock wave from an open end of a duct is shown in Fig. 1. The gas is initially at rest (region 0) and the arriving shock wave produces a steady flow (region 1) in which the pressure p is higher than in region 0. (Here and in the following, p always denotes absolute pressure.) Outflow at this higher pressure cannot be maintained if the flow velocity is subsonic because an expansion wave originates at the exit and travels back into the duct. This reflected wave accelerates the flow further and reduces the pressure to its initial value p_0 at which level steady outflow can take place (region 2).

It is customarily assumed that the readjustment of the exit pressure from p_1 to p_0 takes place instantaneously. All characteristics of the reflected expansion wave originate then at the same point resulting in a "centered" expansion wave as shown in Fig. 1. Actually, because the readjustment of the exit pressure requires a certain time, the reflected wave should appear in the wave diagram as indicated in Fig. 2. The "head" of the expansion wave still originates at

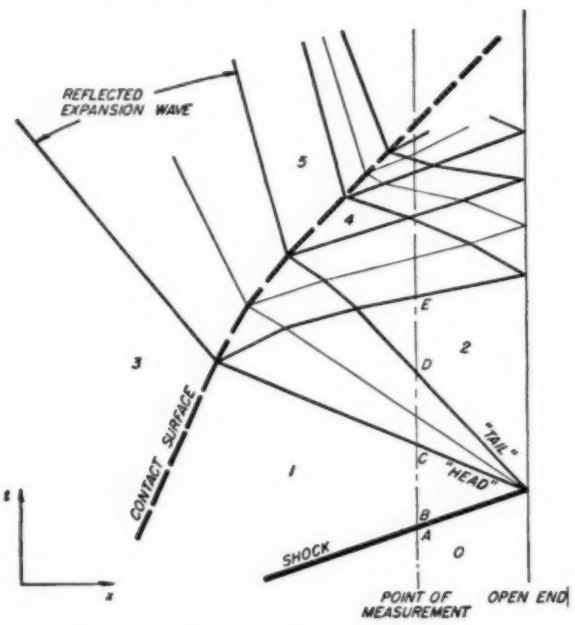


FIG. 1. Wave diagram for a shock wave reflected from an open end. Lag in establishing the steady-flow boundary conditions is neglected and the reflected wave is therefore a centered expansion wave. Interaction of this wave with the contact surface that follows the shock in a shock tube is also indicated.

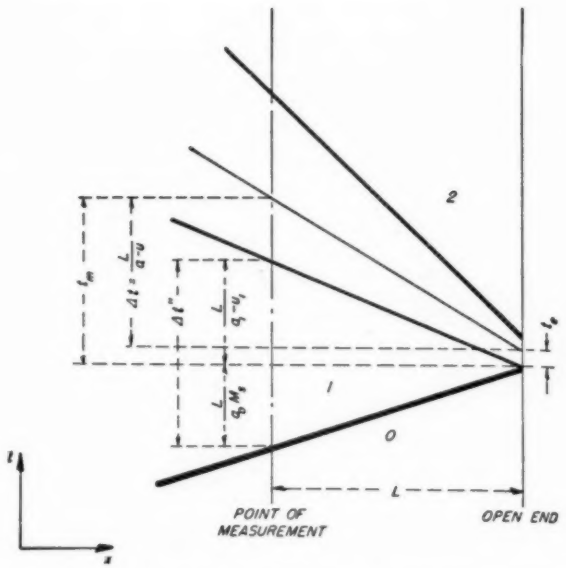


FIG. 2. Wave diagram for a shock wave reflected from an open end when lag is taken into consideration.

[7] J. B. Keller, J. Appl. Phys. **23**, 1267–1268 (1952).
[8] C. C. Hudson, Phys. Rev. **92**(A), 1079 (1953). I am indebted to this author for making the text of his paper available to me.
[9] F. K. Elder and N. de Haas, J. Appl. Phys. **23**, 1065–1069 (1952).

the instant when the shock reaches the open end but the "lag" of later characteristics increases as the exit pressure gradually decays to its steady-flow, equilibrium level. The flow conditions in region 2 are independent of the manner in which the adjustment of the exit pressure takes place. Consequently, the "tail" of the expansion wave has the same slope in the wave diagram, corresponding to the velocity u_2-a_2, in Figs. 1 and 2. Since the characteristics of an expansion wave spread out continually, it is seen that the significance of the lag in establishing the steady-flow boundary conditions diminishes with increasing distance from the exit.

It should also be noted that with increasing strength of the shock wave, the absolute value of the velocity of the "head" of the expansion wave ($|u_1-a_1|$) decreases and, because of the increase of $p_1/p_2=p_1/p_0$, the spread of the reflected expansion wave increases. Both effects combine to lengthen the time required for the entire expansion wave to pass a given location in the duct. The effects of "lag" can, therefore, be expected to loose significance in the case of stronger shocks.

The flow in region 1 (see Figs. 1 and 2) is completely determined by the strength of the arriving shock wave. The flow conditions inside the expansion wave can be computed with the aid of the characteristic relations if a single variable, for instance the pressure, can be measured. It is then also possible to derive the effective flow conditions at the exit from measurements taken at some distance from the exit. With reference to Eqs. (1) and (2), and Fig. 2, it is seen that $Q=$const along any characteristic of the expansion wave as a property of the characteristic while $P=P_1=$const because the flow in region 1 is steady. Since P and Q completely determine the flow conditions, the latter are constant along the characteristics of the expansion wave which are thus straight lines.[10] Any pressure p that occurs at the exit appears, therefore, at a distance L from the exit after a time

$$\Delta t=L/(a-u). \tag{3}$$

This time interval is a function of the pressure and can be computed from the characteristic relation

$$2a/(\gamma-1)+u=P_1, \tag{4}$$

and the condition of isentropic changes of state

$$a/a_1=(p/p_1)^{(\gamma-1)/2\gamma}. \tag{5}$$

Elimination of u from Eqs. (3) and (4), followed by substitution for a from Eq. (5) yields

$$\Delta t(p)=L\left[a_1\frac{\gamma+1}{\gamma-1}\left(\frac{p}{p_1}\right)^{(\gamma-1)/2\gamma}-P_1\right]^{-1}, \tag{6}$$

where a_1, p_1, and P_1 are determined by the strength of the arriving shock wave. From a pressure record taken at a distance L from the open end, the corresponding effective pressure at the open end can now be derived. This procedure involves the time coordinates at the exit and at the point of measurement corresponding to the same value of the pressure and these will be denoted by t_e and t_m, respectively. If time is measured from the the instant of arrival of the shock wave at the exit of the duct, the value of t_e becomes also equal to the lag time during which the exit pressure has adjusted itself from p_1 to p. It is seen from Fig. 2 that

$$t_e(p)=t_m(p)-\Delta t(p). \tag{7}$$

When the shock waves considered become stronger or the measurements are made further away from the exit, both $t_m(p)$ and $\Delta t(p)$ increase and, eventually, become so large compared to $t_e(p)$ that the latter can no longer be accurately determined. This is equivalent to the above statements that the effects of lag become insignificant if the shock wave is too strong or the point under consideration is located too far from the exit. Equations (6) and (7) may also be used to "transfer" a theoretically derived pressure history at the open end to some other station of the duct where it can be compared with experimental results.

In the following, the subscripts of t will be omitted because there is never any doubt as to the meaning of the time coordinate except for this procedure of transferring data from one station of the duct to another.

III. ACOUSTIC THEORY OF SHOCK REFLECTION

The reflection of a sound wave of given frequency from an end of a duct is governed by the specific acoustic impedance. The latter depends on the end configuration and has been determined as a function of frequency for various cases of interest, e.g., reference 11. If the arriving wave has a complicated wave form or is a nonperiodic transient, the required procedure is to obtain first its frequency spectrum (amplitude and phase) by expressing the wave as a Fourier integral; for every frequency component, the amplitude and phase of the reflected component can then be determined from the corresponding impedance and, finally, all reflected components must be superposed to form the reflected wave.

Acoustic theory differs from the theory of large-amplitude waves in two important respects, namely, that pressure changes produced by waves can be added linearly, and that the shape of a propagating wave remains constant. Since waves that propagate out of the duct into the external region rapidly decrease in amplitude, one would expect that most of the errors caused by applying acoustic theory to waves of large amplitude come from the flow conditions inside the duct. It is, therefore, important that the foregoing steps determine the reflected wave from the arriving wave at the *fixed* exit section so that no wave travel along

[10] Note that these statements are valid only for isentropic flow and that they would not hold if wall friction were not neglected.

[11] P. M. Morse, *Vibration and Sound* (McGraw-Hill Book Company, Inc., New York, 1948), second edition, Chaps. VI and VII.

the duct is involved. The propagation of the reflected wave into the duct is treated according to the method outlined in the preceding section so that its subsequent change of wave form is fully accounted for.

This leaves linear superposition of waves as the main limitation of the acoustic theory that must be established by comparison of the results with experimental data.

The Fourier integral of the arriving wave can be expressed as

$$p-p_0=\int_{-\infty}^{+\infty} A(\omega)e^{i\omega(x-a_0t)/a_0}d\omega \tag{8}$$

where ω is the circular frequency, and the complex amplitude of the frequency components is given by

$$A(\omega)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}[p(t)-p_0]e^{-i\omega(x-a_0t)/a_0}dt. \tag{9}$$

Since the problem of wave reflection involves only a fixed value of the x-coordinate, one may set $x=0$ without loss of generality.

The shock wave arriving at the exit produces a sudden increase of the pressure from p_0 to p_1. Equations (8) and (9) must, therefore, be evaluated for a step function. The latter does not satisfy the condition necessary for the existence of the Fourier integral, namely,

$$\int_{-\infty}^{+\infty}[p(t)-p_0]dt=\text{finite}, \tag{10}$$

but this difficulty is overcome by the artifice[12] of setting

$$\left.\begin{aligned} p(t)-p_0&=0 \quad \text{for} \quad t<0\\ p(t)-p_0&=e^{-\beta t} \quad \text{for} \quad t>0.\end{aligned}\right\} \tag{11}$$

This modified step function satisfies condition (10) for any positive value of β, however small. After completion of the entire analysis, one can then let β approach zero.

If conditions (11) are substituted into Eq. (9), one obtains the complex amplitude of the arriving wave as

$$A(\omega)=-\frac{p_1-p_0}{2\pi i}\frac{1}{\omega+i\beta}. \tag{12}$$

The frequency spectrum of the reflected wave will be denoted by $A'(\omega)$. Morse[13] expresses the relation between $A(\omega)$ and $A'(\omega)$ with the aid of a function ψ, defined by

$$A'(\omega)=-A(\omega)e^{-2\psi} \tag{13}$$

and proceeds to show that ψ is related to the specific acoustic impedance z by

$$z=\rho_0 a_0\tanh\psi=\rho_0 a_0(\theta-i\chi) \tag{14}$$

where ρ_0 is the density of the undisturbed gas, and θ and χ are the real and imaginary parts, respectively, of the specific acoustic impedance ratio $z/\rho_0 a_0$.

For the special case of an open-ended circular duct of diameter D fitted with an infinite baffle, the specific acoustic impedance ratio for the plane mode of the oscillations is given by[14]

$$\theta=1-(2/w)J_1(w), \tag{15}$$

$$\chi=\frac{4}{\pi}\int_0^{\pi/2}\sin(w\cos\varphi)\sin^2\varphi d\varphi, \tag{16}$$

where

$$w=\omega D/a_0 \tag{17}$$

and $J_1(w)$ is the Bessel function of the first kind and first order. The integral in Eq. (16) is closely related to the Struve function $\mathbf{S}_1(w)$. It is easily verified[15] that

$$\chi=(2/w)\mathbf{S}_1(w). \tag{18}$$

The asymptotic values of θ and χ for small and large arguments will also be required. They are given by[14]

$$\left.\begin{aligned} &\theta\to w^2/s, \quad \chi\to 4w/(3\pi) \quad \text{for} \quad w\to 0\\ &\theta\to 1, \quad\quad \chi\to 4/(\pi w) \quad \text{for} \quad w\to\infty\end{aligned}\right\}. \tag{19}$$

It is now possible to set up the Fourier integral for the reflected wave. The exit pressure following the arrival of the shock wave is given by

$$p-p_0=(p_1-p_0)+\int_{-\infty}^{+\infty}A'(\omega)e^{-i\omega t}d\omega. \tag{20}$$

Equation (14) may also be written in the form

$$\tanh\psi=(e^{2\psi}-1)/(e^{2\psi}+1)=\theta-i\chi$$

from which one obtains

$$e^{-2\psi}=(1-\theta+i\chi)/(1+\theta-i\chi). \tag{21}$$

This and Eq. (12) must be substituted into Eq. (13) to obtain $A'(\omega)$. The integration variable in Eq. (20) must be changed to w since θ and χ are functions of this variable and not of ω. Using the abbreviation

$$\beta'=\beta D/a_0 \tag{22}$$

and the nondimensional time

$$\tau=a_0 t/D, \tag{23}$$

one obtains the solution of the problem in the form of a

[12] E. A. Guillemin, *The Mathematics of Circuit Analysis* (John Wiley and Sons, Inc., New York, 1949), p. 532.

[13] See reference 11, p. 239.

[14] See reference 11, p. 333. The notation θ_0 and χ_0 is used there to distinguish the plane mode from higher modes. Since the average of the latter over the cross section of the duct is zero, these do not contribute to the effective, one-dimensional exit condutions required here and, thus, need not be considered.

[15] E. Jahnke and F. Emde, *Tables of Functions* (Dover Publications, Inc., New York, 1945), fourth edition, p. 211.

Fourier integral

$$I=\frac{p-p_0}{p_1-p_0}=1+\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{(1-\theta+i\chi)e^{-iw\tau}}{(1+\theta-i\chi)(w+i\beta')}dw. \quad (24)$$

The pressure ratio I indicates the deviation of the exit pressure from its steady-flow equilibrium value p_0. It is expressed in terms of the pressure rise through the incident shock wave but there is no further dependence of I on the shock strength since the right-hand side of Eq. (24) is a function of τ only. The solution is also independent of the properties of the gas involved. However, the latter affect the actual lag times through the value of the speed of sound which is used to make the time coordinate nondimensional.

The nature of the solution is analyzed in Appendix I where it is shown that I is made up of an infinite number of damped harmonic oscillations superposed on a dominant exponential decay. Actual values of I were obtained by numerical integration (see Appendix II) and the final results are shown in Table I and Fig. 8.

TABLE I. Result of Numerical Integration of Eq. (24) ($\lambda=14.5$, see Appendix II).

τ	I_λ	R_λ	$I=I_\lambda+R_\lambda$
0	0.986	0.014	1.000
0.05	0.964	0.002	0.966
0.25	0.834	0	0.834
0.50	0.643	−0.001	0.642
1.00	0.319	−0.001	0.318
1.50	0.147	−0.001	0.146
2.00	0.070	0	0.070
3.00	0.016	−0.002	0.014

IV. EXPERIMENTAL TECHNIQUES

A shock tube, e.g., references 4, 16–18 is eminently suitable to carry out experiments of the kind required here. If the end of the low-pressure chamber is open to the atmosphere and a pressure transducer is mounted several diameters from the exit, passage of the shock and of the reflected expansion wave can be recorded. The latter is again reflected from the contact surface that follows the shock wave—as indicated in Fig. 1—and the duct must be long enough to ensure that the entire expansion wave has passed the pressure transducer before the first reflection from the contact surface arrives there. This requires a minimum length of the expansion chamber which can be found (for any selected shock strength and location of the point of measurement) from the simple wave diagram for a shock tube. From the same wave diagram, one can then also determine the required minimum length of the high-pressure chamber: The expansion wave which propagates into the latter, following the breaking of the shock tube diaphragm, is reflected from the closed end of that chamber and must then not reach the pressure transducer before the transient that is to be recorded is completed.

The shock tube was constructed of sections of hard-drawn, seamless brass tubes which had an outer diameter of 3.50 inches and an inner diameter of 3.23 inches. The low-pressure chamber was made up of several sections combining to a total length of 16 feet and by rearranging these, the pressure transducer could be placed at various distances from the tube exit. The high-pressure chamber was 12 feet long. Each chamber was provided with a flange and these could be held together by means of a threaded coupling which permitted quick changes of the diaphragm. The latter consisted of photographic film and, depending on the pressures required, between one and three sheets were used. These were tightly held by means of two O-rings set into the flanges. The high-pressure chamber was filled with compressed air, and a solenoid-operated plunger was used to rupture the diaphragms. The end of the low-pressure chamber was open and the end configuration could be varied between a sharp-edged exit and a rigid baffle of 2 feet in diameter.

The pressure was measured by means of a condenser-type transducer combined with the electronic signal converter supplied by the manufacturer.[19] The pressure-sensitive diaphragm of this instrument has a diameter of about $\frac{3}{16}$ inches and a diaphragm rated at 50 psig was used. The frequency response of the system is stated by the manufacturer to be flat from zero to one-third of the natural diaphragm frequency and to be off from 3 to 5% at two-thirds of the natural frequency. Passage of a shock wave excited the resonance oscillations of the diaphragm and their frequency was estimated to be in the vicinity of 20 000 cps. Judging from this and the appearance of the records, the frequency response of the system seemed to be adequate for the planned experiments. The resonance oscillations did not introduce any difficulties for the evaluation of the records since they decayed almost completely before the reflected expansion wave arrived at the pickup. The electronic unit provides for compensation of the basic nonlinearity of a condenser-type transducer. When properly adjusted, it was found that the calibration could be taken as linear over the required pressure range.

The electric signal from the pressure measuring equipment was connected to one channel of a dual-beam, cathode-ray oscilloscope and could be recorded photographically by means of an oscillograph record camera.[20] The second channel of the oscilloscope was connected to a frequency standard (2500 cps) to provide a

[16] Bleakney, Weimer, and Fletcher, Rev. Sci. Instr. **20**, 807–815 (1949).

[17] J. Lukasiewicz, "Shock tube theory and applications," Natl. Research Council Can., Rept. MT-10, 18 January 1950.

[18] Glass, Martin, and Patterson, "A theoretical and experimental study of the shock tube," University of Toronto, Institute of Aerophysics, UTIA Rept. No. 2, November 1953.

[19] Pressure pickup Model HR-3 and electronic indicator type ST-12 made by the Rutishauser Corporation, Pasadena, California.

[20] DuMont dual-beam cathode-ray oscilloscope type 322-A combined with a Fairchild Oscillo-Record Camera.

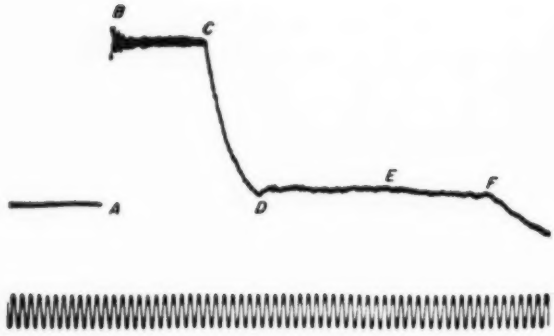


FIG. 3. Typical pressure record obtained on a continuously moving film for a shock pressure ratio of 1.56. The pressure is measured 30 inches from the duct exit, and the frequency of the reference sine wave is 2500 cps. The letters *A*, *B*, *C*, *D*, and *E* correspond to the similarly labelled points in Fig. 1.

calibration of the time scale. A typical record is reproduced in Fig. 3. This record shows the pressure history in agreement with that derived from a wave diagram (Fig. 1), and corresponding points are labeled *A*, *B*, *C*, etc. The shock wave passes the pressure transducer at *A* and raises the pressure, *B*. There follows an interval of constant pressure until the head of the reflected expansion wave arrives, *C*. During passage of the expansion wave, the pressure decays to almost atmospheric level,[21] *D*, and remains then constant until the reflection of the expansion wave from the contact surface arrives, *E*. This produces a gradual pressure drop that is so small that it is hardly noticeable on the record. Finally, the expansion wave that is produced by the rupture of the shock-tube diaphragm (not shown in Fig. 1) arrives, *F*, and reduces the pressure to well below atmospheric. The pressure between *D* and *E* is slightly higher than atmospheric since wall friction causes a small pressure drop in the flow between the point of measurement and the exit of the duct. Records of this kind yield the entire pressure history, and since the film is moving continuously, there are no problems of synchronization. However, since the maximum film speed of the available camera was only 5 feet per second, the time resolution in the region of the reflected expansion wave was not adequate except in cases of stronger shock waves (see Sec. II). It was, therefore, decided to use a single-sweep technique to increase the time resolution. A schematic diagram of the setup and instrumentation is shown in Fig. 4.

[21] It may be noted that the expansion wave is, apparently, followed at *D* by a weak compression wave. This feature was noted on practically all records and was also obtained with other types of pressure transducers. Its explanation is not clear. It could represent a gas oscillation across the duct or a compression wave that reenters the duct from the exterior. Such waves were observed by Glass (see reference 4); since he used schlieren techniques which respond to the derivative of the density, his records may show strong signals while the actual magnitude of the pressure variations may be quite small. No further investigation of this phenomenon was undertaken since no cases were observed where it reached a significant magnitude. Some recent work (to be published) seems to indicate that this detail of the records is related to the growth and subsequent downstream motion of the vortex formed at the duct exit after arrival of the shock wave.

The shock tube and auxiliary equipment are indicated at the top of the figure. The electric signal from the pressure measuring equipment is connected to the input of one channel of the oscilloscope (Y_1). The same signal is also fed to a simple pulse generator consisting of an amplifier with a type 2050 thyratron output stage. Passage of the shock wave at the pressure transducer produces thus a single pulse which is used to trigger a single sweep of the beam of the oscilloscope. Immediately before breaking the shock tube diaphragm, an extra sweep can be produced by means of a manually operated pulse generator (discharge of a condenser). This sweep serves as base line of the records, corresponding to atmospheric pressure. Timing calibration is obtained by means of beam intensity (Z) modulation. The output from a 2500 cps frequency standard is connected to the input of the second channel of the oscilloscope (Y_2) and the horizontal sweep of the latter is synchronized with a convenient multiple of the standard frequency as determined by the stationary pattern on the oscilloscope screen. An amplified signal is then taken directly from one of the horizontal deflection plates (X_2) and connected to the Z-terminal of the first channel. The advantages of this scheme are that various frequencies can be obtained from a fixed-frequency standard, and also that the triangular wave form of the voltage taken at the X-deflection plate yields somewhat better defined timing marks than a sinusoidal wave.

The oscilloscope traces were recorded photographically on a stationary film. If sweep speed and position were properly adjusted, any desired portion of the pressure transient could be recorded. This adjustment had to be done by trial and error involving each time taking and developing a photograph. The use of an oscilloscope camera[22] based on the Polaroid-Land process greatly helped to reduce the time required for adjustments. The final records were then taken on regular film and evaluated with the aid of a telereader. A sample record

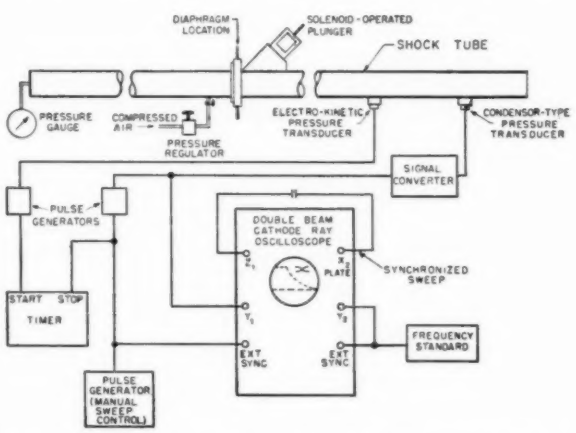


FIG. 4. Schematic diagram of shock tube, auxiliary equipment, and instrumentation for the single-sweep recording technique.

[22] DuMont oscillograph record camera type 297.

obtained in this manner, using a calibration frequency of 10 000 cps is reproduced as Fig. 5. This record shows only the final part of the interval of constant elevated pressure following the shock wave and the pressure drop through the reflected expansion wave. The final pressure is again a little higher than atmospheric because of the effects of wall friction.

Because of the linearity of calibration of the pressure instrumentation, a single reference value is sufficient to evaluate the records. This is obtained by determining the pressure rise through the shock wave from the shock velocity. A second pressure transducer was mounted on the shock tube (see Fig. 4) so that, depending on the arrangement used, the distance between the two pickups was either 2.0 or 2.5 feet. The second transducer was of the electro-kinetic type[23] which delivers an electric output signal without any auxiliary equipment to a pulse generator identical to the one already mentioned. The outputs of the two pulse generators are used to "start" and "stop" an electronic counter. The travel time of the shock wave between the two pickups can thus be measured with an uncertainty of at most one count (10 microseconds). In a preliminary experiment, both transducers were mounted at the same position along the shock tube and it was verified that the counter reading under those conditions never exceeded one count. Let $\Delta t'$ denote the time in which the shock wave travels the distance L' between the pressure pickups. From

$$\Delta t' = L'/a_0 M_S \tag{25}$$

one obtains then the shock Mach number M_S which, in turn, is related to the pressure rise across the shock wave through the Rankine-Hugoniot shock relation[24]

$$p_1 - p_0 = 2\gamma p_0 (M_S^2 - 1)/(\gamma + 1). \tag{26}$$

This method of calibration of the pressure measuring equipment is not only convenient but provides also a check of the calibration for each individual record. Since it is, essentially, based on the deviation of the shock velocity from the speed of sound, it becomes

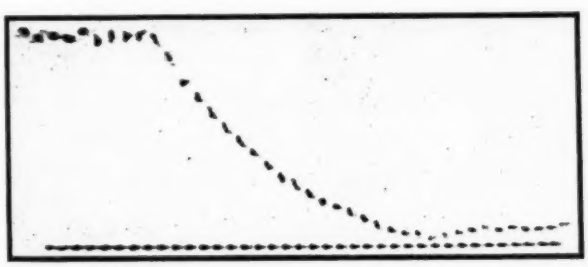


FIG. 5. Typical pressure record obtained by the single-sweep technique for a shock pressure ratio of 1.56. Pressure is measured at 12 inches from the duct exit and the frequency of the timing marks is 10 000 cps.

[23] E. V. Hardway, Jr., Instruments **26**, 1186–1188 (1953). Manufactured by the Beta Corporation, Richmond, Virginia.

[24] See, for example, reference 1, Sec. 67.

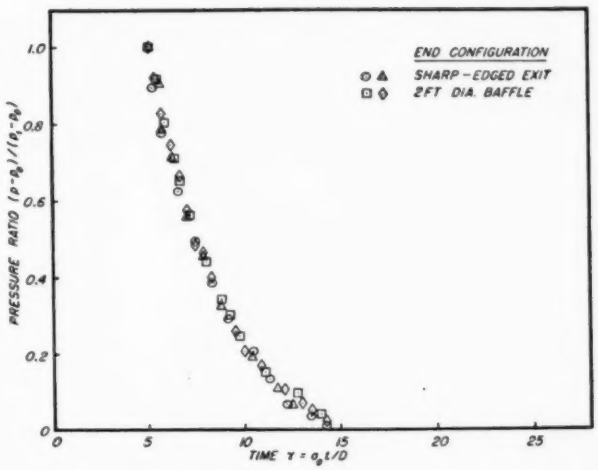


FIG. 6. Typical plot of measured pressure ratios indicating that the results are well reproducible and not affected by the exit configuration ($p_1/p_0 = 1.56$).

inaccurate, for a given accuracy of the time measurement, if M_S approaches unity to closely. It can be verified, however, that for $p_1/p_0 > 1.5$ and $L' = 2.5$ feet, a timing error of 10 microseconds (one count) determines the pressure rise through the shock wave with an accuracy better than 3.6%. The uncertainty of the absolute pressure is then about 1%.

An additional check on this method of calibration was obtained by comparing the results with two other methods of calibration. In one of these, the instruments are calibrated statically. In the other method of comparison, the time between passage of the shock wave at the pressure transduces and arrival of the head of the reflected expansion wave is obtained from oscilloscope records. This interval $\Delta t''$ is given by (see Fig. 2)

$$\Delta t'' = \frac{L}{a_0}\left(\frac{1}{M_S} + \frac{a_0}{a_1 - u_1}\right), \tag{27}$$

where M_S, a_1/a_0, and u_1/a_0 are again related to the pressure rise across the shock wave through the Rankine-Hugoniot equations.

These three methods of calibration were compared in a series of 19 experiments in which the shock pressure ratio was varied between 1.56 and 2.40. In 11 of these experiments, the pressure rise across the shock wave, computed on the basis of the shock velocity deviated from the mean of all three methods by 1.8% or less, and in one case only, did the discrepancy exceed 3.6%. On the basis of these results, the described dynamic method of calibration was considered to be satisfactory.

V. RESULTS AND DISCUSSION

Since the significance of "lag" in establishing the steady-flow exit conditions diminishes with increasing distance from the exit while the effects of wall friction increase, all measurements should be carried out as close as possible to the exit of the duct. At the same

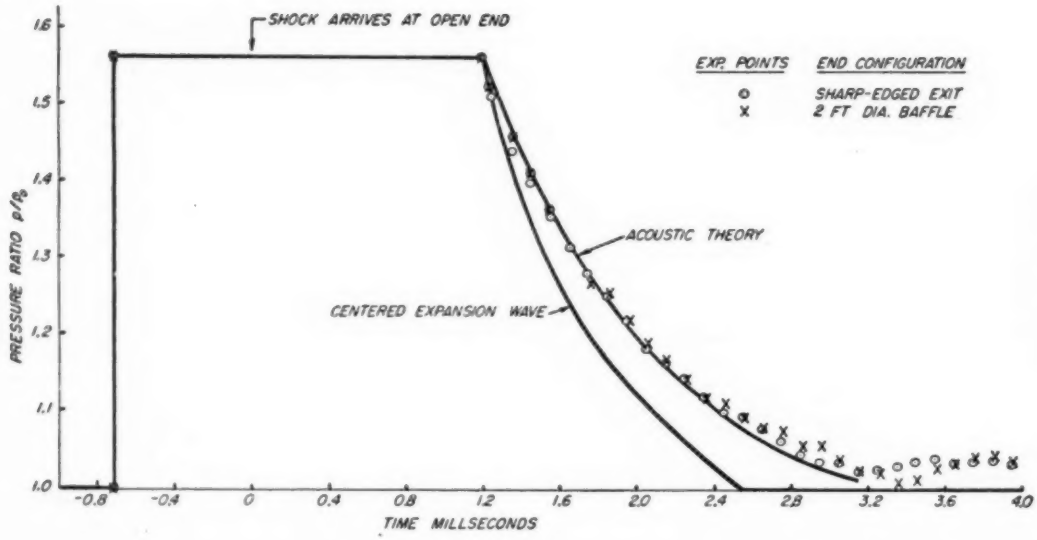


Fig. 7. Comparison of experimental results, centered expansion wave, and acoustic theory at the point of measurement (12 inches from the duct exit) for a shock pressure ratio of 1.56.

time, one must stay far enough from the exit to have substantially one-dimensional flow. A distance of one foot was considered a suitable compromise. A few measurements were also taken at 2.5 feet (see, for example, Fig. 3), and the results were in qualitative agreement with those obtained at the smaller distance although, for the aforementioned reasons, a quantitative evaluation of the results was no longer considered to be reliable.

All data were well reproducible. This is illustrated in Fig. 6 which shows the results of several typical experiments. The exit of the duct was either sharp-edged or provided with a baffle of 2 feet in diameter. In spite of these extreme variations of the end configuration, no measurable effect on the process of wave reflection could be detected inside the duct. This is not surprising since theoretical investigations indicate[14] that the actual baffle shape is of little importance as far as the phenomena *inside* the duct are concerned. This finding simplified the experimental work since only one end configuration had to be investigated.

The experimental results must be compared with theory based either on the conventional steady-flow boundary conditions (centered expansion wave) or on the acoustic theory, and as pointed out in Sec. II, such a comparison may be made either at the point of measurement or at the exit.

A comparison at the point of pressure measurement is presented in Fig. 7 for a shock pressure ratio of 1.56. It may be seen that acoustic theory and experimental results for the two extreme exit configurations are in good agreement while the pressure decay predicted on the basis of the steady-flow boundary conditions is considerably steeper. In these experiments, the tail of the expansion wave lags by about one-half millisecond.

Comparison between theory and experiment at the duct exit is shown in Figs. 8(a), (b), and (c) for shock pressure ratios of 1.56, 1.74, and 1.93, respectively. The plotted points represent the mean of 10 experiments except in Fig. 8(c) where only four experiments were used. The estimated uncertainty based on the scatter of individual experiments is indicated by the shaded bands and their marked widening with increasing shock strength may be noted. This again is the result of the reduced significance of lag in the case of stronger shock waves: The magnitude of the effect to be measured

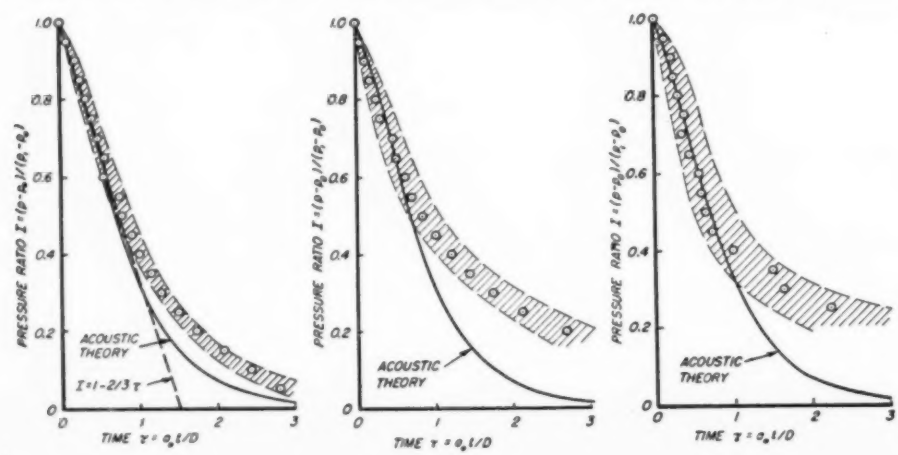


FIG. 8. Decay of the exit pressure as predicted by acoustic theory compared to data derived from measurements taken at 12 inches from the exit. The shaded bands indicate the estimated uncertainty based on the scatter of individual experiments: (a) $p_1/p_0=1.56$, (b) $p_1/p_0=1.74$, and (c) $p_1/p_0=1.93$.

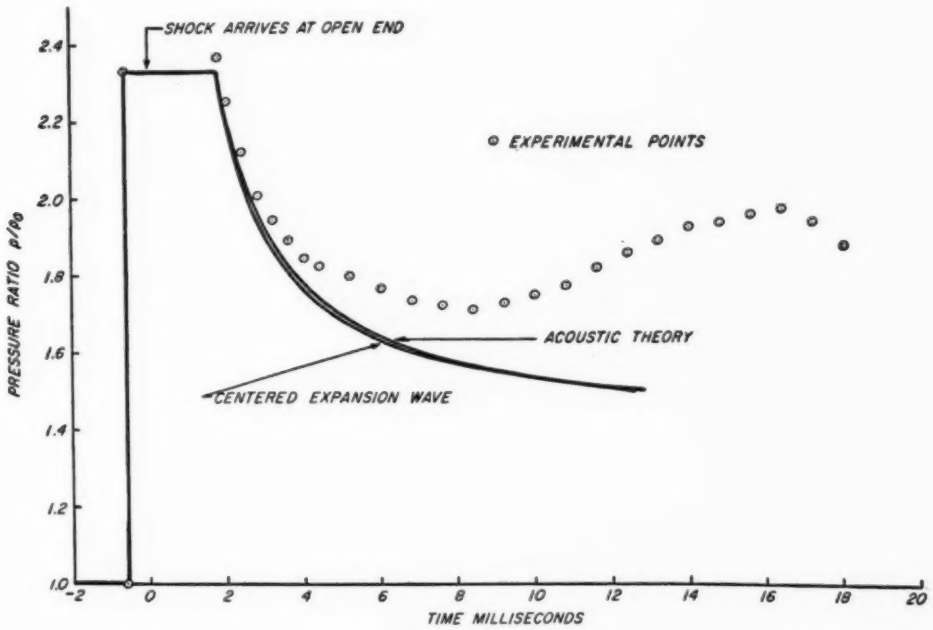


FIG. 9. Comparison of experimental results, centered expansion wave, and acoustic theory at the point of measurement (12 inches from the duct exit) for a shock pressure ratio of 2.33.

approaches that of the experimental errors. Since the results of the acoustic theory are independent of shock strength, the corresponding curves, plotted from Table I, are the same in the three figures. In these plots, the steady-flow boundary conditions are represented by the instantaneous drop of the pressure ratio from 1.0 to zero.

The satisfactory agreement between acoustic theory and experimental results can again be noted, particularly during the early and most significant portion of the pressure decay. Since the flow velocity at the tail of the expansion wave is about twice that at the head, the friction effects at the former are increased by a factor of about four. This explains qualitatively why the agreement between theory and experiment is so much better during the early part of the decay curve. The experimental results seem to indicate (see Figs. 3 and 5) that the exit pressure decays to a value above atmospheric where it remains practically constant. Since the flow has then become steady, the exit pressure must have become atmospheric which is the steady-flow boundary condition. However, the transfer of the experimental data to the exit, according to Sec. II, does not take friction into account and leads, therefore, to an equilibrium exit pressure that is too high. If friction effects in the nonsteady flow could be taken into consideration, the experimental data for τ greater than about 3 would have to agree with theory, and the location of points for smaller values of τ would also be improved.

The reflected expansion wave can accelerate the flow at the exit, at most, to sonic velocity. In air ($\gamma=1.4$), this occurs at a shock pressure ratio of about 1.94. For still stronger shocks, outflow remains sonic but the exit pressure stabilizes at a higher-than-atmospheric level. Under these conditions, the pressure variations inside the duct become so slow that any lag at the exit is quite insignificant. This is illustrated in Fig. 9 where theory and experiment are again compared at the point of measurement for a shock pressure ratio of 2.33. It is seen that the theoretical results based on the acoustic theory and the steady-flow boundary conditions, respectively, practically coincide while the deviations of the experimental results are caused by friction. One is thus led to the conclusion that the acoustic theory adequately describes the pressure decay at the exit as long as the outflow produced is subsonic, while for shocks strong enough to produce sonic outflow, the lag effects become so small that they need not be considered.

A comparison of the experimental results in Fig. 9 with those shown in Figs. 3 and 7 reveals certain modifications of the pressure history. After passage of the shock wave, the pressure no longer remains constant but continues to rise. This phenomenon has also been observed elsewhere[25,26]; it is apparently caused by friction but has not yet been completely explained. Another interesting feature is the marked pressure rise after an initial decay. This effect is produced when the flow behind the contact surface becomes supersonic.[27] With reference to Fig. 1, the reflected expansion wave, after crossing the contact surface, accelerates the flow in region 3 until the higher velocity in region 5 is reached. Figure 1 is drawn for the case that the flow in region 5 is subsonic but if the waves become stronger, the flow

[25] Emrich, Mack, and Shunk, Phys. Rev. **97** (A), 375 (1953).
[26] J. E. Mack, "Density measurement in shock tube flow with the chrono-interferometer," thesis, Lehigh University, Institute of Research, Tech. Rept. No. 4, April 15, 1954.
[27] At first, this effect was ascribed to a constricting action of the growing boundary layer until the above explanation became apparent.

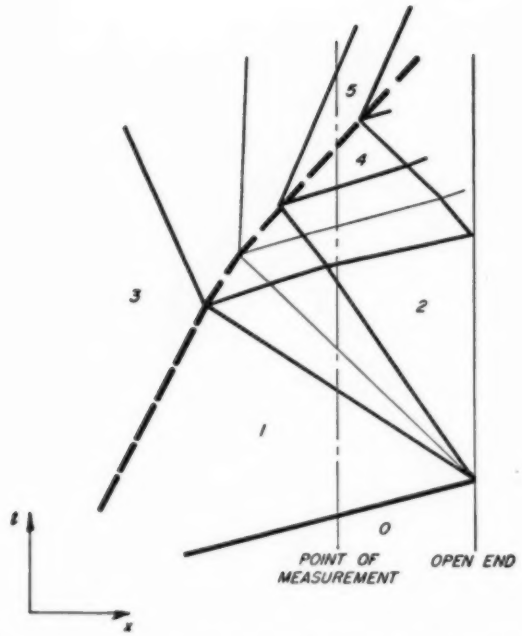


FIG. 10. Modification of the wave diagram in Fig. 1 when the expansion wave accelerates the flow behind the contact surface to a supersonic velocity (region 5).

in region 5 may become supersonic (while the same flow velocity in region 4 on the other side of the contact surface still remains subsonic). The correspondingly modified wave diagram is shown in Fig. 10. It is seen that some of the characteristics of the expansion wave, after crossing the contact surface, are swept downstream and therefore pass the location of the pressure transducer for a second time but now in reverse order. Consequently, the measured pressure must rise again after an initial drop. Under the conditions under which the shock tube experiments were carried out (air at room temperature used as driver gas), the flow in region 5 would become supersonic for shock pressure ratios exceeding about 1.75.

Both the acoustic theory and the manner in which the experiments are evaluated yield the average pressure for the exit section of the duct. Actually, when a shock wave of infinitesimal strength reaches the exit, the shock induced flow at the center of the exit section remains unaffected until a disturbance wave arrives there after a time $D/2a_0$. When stronger shocks are considered, the disturbance wave propagates with the speed a_1 in a fluid flowing with the velocity u_1 and the center of the exit section is reached after a time $D/2(a_1^2-u_1^2)^{\frac{1}{2}}$. Twice this time rather than D/a_0 should, therefore be used to make the lag times nondimensional.[28] Thus the nondimensional time introduced in the acoustic theory should be interpreted as

$$\tau=\frac{(a_1^2-u_1^2)^{\frac{1}{2}}}{a_0}\frac{a_0 t}{D}. \tag{28}$$

[28] This was already pointed out in reference 5.

For shocks of infinitesimal strength, Eqs. (28) and (23) become identical, and it is easily verified that for shock pressure ratios less than 2.0 (for $\gamma=1.4$), the correction factor $(a_1^2-u_1^2)^{\frac{1}{2}}/a_0$ deviates from unity by less than 3% which represents a negligible error of the lag time. For $p_1/p_0=2.5$, the actual lag times are 9% longer than those computed without taking this effect into account. However, for shocks of such strength, the lag effects are already so unimportant that this error again becomes insignificant. It seems therefore that it should be entirely adequate to set the correction factor in Eq. (28) equal to 1.0.

The consequences of the lag in establishing the steady-flow boundary conditions are clearly noticeable in Fig. 7. The later portions of the reflected expansion wave are delayed by a time during which a sound wave can travel two to three times across the exit section of the duct. Lags of this magnitude may become significant if, for instance, a critical event is to take place as soon as the pressure has dropped to a prescribed value. As a typical example of this kind, consider inflow of fresh air into the combustion chamber of a pulse-jet which starts as soon as the pressure has dropped to the level of the stagnation pressure at the engine inlet. If the ratio of engine length to diameter is sufficiently small, the results of a wave diagram analysis might be significantly affected by lag.

It is also seen in Fig. 7 that the actual rate of pressure decay at the head of the reflected expansion wave is considerably lower than that predicted on the basis of the steady-flow boundary conditions. This means that the computed acceleration of a gas particle, as it enters the expansion wave, would be in error if lag were neglected. The acceleration of a gas particle at the head of the reflected expansion wave and distance L from the exit is given by (see Appendix III)

$$\frac{Du}{Dt}=\frac{2a_0a_1^2(a_1-u_1)(p_1-p_0)}{(\gamma+1)a_0a_1(p_1-p_0)L+3\gamma(a_1-u_1)^2p_1D}. \tag{29}$$

The corresponding acceleration in the absence of lag is obtained from Eq. (29) by setting $D=0$. The ratio of the actual acceleration to that computed on the basis of the steady-flow boundary conditions is thus given by

$$r=\left(\frac{Du}{Dt}\right)\Big/\left(\frac{Du}{Dt}\right)_{D=0}=\frac{1}{1+\sigma D/L}, \tag{30}$$

where

$$\sigma=\frac{3\gamma(a_1-u_1)^2p_1}{(\gamma+1)a_0a_1(p_1-p_0)} \tag{31}$$

is a parameter that depends only on the strength of the shock wave and the value of γ. For $\gamma=1.4$, r is plotted in Fig. 11 as function of the shock pressure ratio and for values of $L/D=5$, 10, and 20. The actual acceleration is also plotted in the dimensionless form (D/a_0^2)

$\times(Du/Dt)$ for $L/D=10$. It is seen that r approaches 1.0 only if the shock is sufficiently strong (the effect of friction on the acceleration is not considered here), or if the point under consideration is located sufficiently far from the exit. For example, for $p_1/p_0=1.2$ and $L/D=10$, the conventional theory would yield an acceleration at the head of the reflected expansion wave that is almost twice the correct value. The actual acceleration produced by such a weak wave in a duct of 3 inches diameter and at room temperature would in this case be about 2.2×10^5 ft/sec^2 and thus by no means be small.

Accurate knowledge of the acceleration may be required, for instance, in investigations of the stability of contact surfaces[29] which are already noticeably affected by gravity. Similarly, the stability and rate of heat release of flame fronts is strongly influenced by accelerations[30] although virtually no quantitative data are as yet available. In any wave diagram analysis involving such phenomena, it may become important to determine accelerations more accurately than is possible on the basis of the steady-flow boundary conditions.

On the other hand, in many practical applications of wave diagrams, consideration of lag may not be warranted in view of other errors that may be introduced by some simplifying assumptions about the phenomenon to be studied. The results of this investigation should make it possible to estimate, for any particular problem, whether or not the effects of lag are significant enough to be considered in the analysis. If they are to be included, the presented results of the acoustic theory, or even a linear approximation of the decay curve,[31] should be adequate.[32]

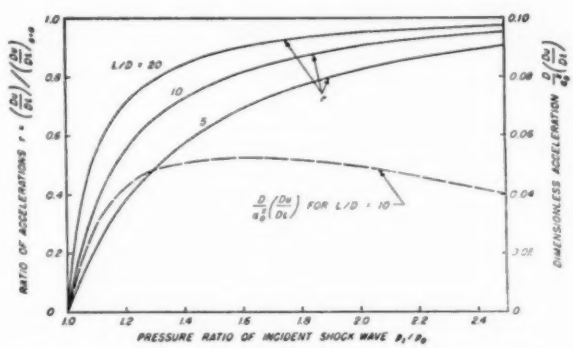


FIG. 11. Variation of the acceleration ratio r and of the (non-dimensional) acceleration at the head of the reflected expansion wave with the strength of the incident shock wave for $\gamma=1.4$.

[29] G. I. Taylor, Proc. Roy. Soc. (London) **A201**, 192 (1950); D. J. Lewis, *ibid.* **A202**, 81 (1950).

[30] G. H. Markstein, *Instability Phenomena in Combustion Waves* in Fourth Symposium (International) on Combustion (Williams and Wilkins, Baltimore, 1953), pp. 44–59.

[31] See reference 2, p. 148.

[32] Wave diagram procedures based on these boundary conditions are considerably more time consuming than those based on a centered expansion wave. It was found recently that the boundary conditions could be satisfactorily approximated by a centered expansion wave provided its center is shifted to a suitably chosen new location. These findings will be published in a later communication.

VI. ACKNOWLEDGMENTS

The author is greatly indebted to Mr. L. M. Somers for conducting the experiments, and Dr. T. R. Goodman and Mr. H. R. Lawrence for helpful discussions of the mathematical problems concerning the acoustic theory. The assistance of Mr. R. Blazer in the evaluation of the photographic records is also acknowledged.

APPENDIX I. NATURE OF THE RESULT REPRESENTED BY EQ. (24)

If one considers w as a complex variable, then I can be expressed in terms of a contour integral over the lower half of the complex plane. This path of integration is required because for the positive values of τ that are of interest the integrand vanishes at infinity only for negative imaginary values of w. Denoting the integrand briefly by $F(w)$, one has, therefore,

$$I=1+\frac{1}{2\pi i}\int_{-\infty}^{+\infty}F(w)dw=1+\frac{1}{2\pi i}\oint F(w)dw$$
$$=1-\Sigma\text{ residues.}\quad\text{(A1)}$$

The negative sign is required here because the contour of the lower half-plane is travelled in clockwise (negative) direction. In order to obtain the residues, one must find the roots of the denominator of $F(w)$. One root is, clearly, given by

$$w=-i\beta'\to0.\quad\text{(A2)}$$

According to Eqs. (19), the corresponding residue becomes equal to 1.0 which cancels the first term on the right-hand side of Eq. (A1).

The other factor of the denominator of $F(w)$ may be expressed as an integral. Since the Bessel function $J_1(w)$ may be written in the form[33]

$$J_1(w)=(2w/\pi)\int_0^{\pi/2}\cos(w\cos\varphi)\sin^2\varphi d\varphi,$$

one obtains the equations for the remaining roots w_n with the aid of this and Eq. (16)

$$1+\theta-i\chi=2-\frac{4}{\pi}\int_0^{\pi/2}e^{iw_n\cos\varphi}\sin^2\varphi d\varphi=0.\quad\text{(A3)}$$

Setting

$$w_n=\Omega_n-i\mu_n,\quad\text{(A4)}$$

with $\mu_n\geqslant0$ because all roots of interest must lie in the lower half-plane, one can separate Eq. (A3) into its real and imaginary parts

$$\int_0^{\pi/2}e^{\mu_n\cos\varphi}\cos(\Omega_n\cos\varphi)\sin^2\varphi d\varphi=\frac{\pi}{2},\quad\text{(A5)}$$

$$\int_0^{\pi/2}e^{\mu_n\cos\varphi}\sin(\Omega_n\cos\varphi)\sin^2\varphi d\varphi=0.\quad\text{(A6)}$$

[33] See reference 15, p. 151.

None of the roots w_n are real because the inequality

$$\int_0^{\pi/2}\cos(\Omega_n\cos\varphi)\sin^2\varphi d\varphi<\int_0^{\pi/2}\sin^2\varphi d\varphi=\frac{\pi}{4}$$

indicates that Eq. (A5) cannot be satisfied for $\mu_n=0$.

Clearly, $\Omega_0=0$ satisfies Eq. (A6) regardless of the corresponding value of μ_0 which is then determined by Eq. (A5). All other roots are complex and occur in pairs $\pm\Omega_n-i\mu_n$. It is evident from the exponential factor in $F(w)$ that residues of each pair of these roots combine to one term representing a damped harmonic oscillation of some amplitude B_n and phase δ_n. Eq. (A1) thus assumes the form

$$I=B_0e^{-\mu_0\tau}+\sum_{n=1}^{\infty}B_ne^{-\mu_n\tau}\sin(\Omega_n\tau+\delta_n). \qquad (A7)$$

Inspection of Eq. (A5) reveals that the smallest value of μ is obtained for $\Omega=0$. For sufficiently large τ, the value of I tends therefore to decay exponentially as indicated by the first term of Eq. (A7). However, while the constants of this term can be determined by means of simple numerical integrations ($\mu_0=1.460$ and $B_0=1.298$), no reasonable way could be found to determine the other roots. Equation (A7) indicates therefore merely the nature of the solution, but for actual use the relation between I and τ must be found by numerical integration of Eq. (24). This is preferable in any case because a large number of terms would probably have to be added up in order to yield a reasonably accurate decay curve for small values of τ which are of particular interest here.

APPENDIX II. NUMERICAL INTEGRATION OF EQ. (24)

In order to bring Eq. (24) into a form that is suitable for numerical integration, it must first be rationalized so that the real and imaginary terms can be separated. It is then found that the real part is an even function of w while the imaginary part is odd (it is seen from Eqs. (15) and (16) that θ is even and χ is odd). The integral of the imaginary terms is thus zero, as must be the case. Some of the real terms include β' as a factor while others do not. Collecting them accordingly, one obtains

$$I=1+\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{2\chi\cos w\tau-(1-\theta^2-\chi^2)\sin w\tau}{[(1+\theta)^2+\chi^2](w^2+\beta'^2)}wdw$$

$$-\frac{\beta'}{2\pi}\int_{-\infty}^{+\infty}\frac{2\chi\sin w\tau+(1-\theta^2-\chi^2)\cos w\tau}{[(1+\theta)^2+\chi^2](w^2+\beta'^2)}dw. \qquad (A8)$$

It is permissible to let β' go to zero in the first integral since the lowest power in the numerator [see Eqs. (19) for the first terms of the series expansion of θ and χ] is w^2 so that the integrand remains finite at $w=0$. The situation is different in the second integral where the lowest power in the numerator is w^0. The integrand has therefore a singularity at $w=0$. Since the integral is multiplied by β' any contribution can only come from a narrow range of integration between $-\epsilon$ and $+\epsilon$ where ϵ is an arbitrarily small but finite number. One obtains then for the last term of Eq. (A8)

$$-\frac{\beta'}{2\pi}\int_{-\epsilon}^{+\epsilon}\frac{dw}{w^2+\beta'^2}=-\frac{1}{2\pi}\int_{-\epsilon/\beta'}^{+\epsilon/\beta'}\frac{dy}{y^2+1}$$

$$=-\frac{1}{\pi}\arctan\frac{\epsilon}{\beta'}\to-\tfrac{1}{2}\quad\text{for}\quad\beta'\to0. \qquad (A9)$$

The final result which is suitable for numerical integration for selected values of τ is thus given by

$$I=\tfrac{1}{2}+\frac{1}{\pi}\int_0^{\infty}\frac{2\chi\cos w\tau-(1-\theta^2-\chi^2)\sin w\tau}{[(1+\theta)^2+\chi^2]w}dw \qquad (A10)$$

where use has been made of the fact that the integrand is an even function so that the lower limit of the integration can be changed from $-\infty$ to zero provided the integral is, at the same time, multiplied by 2.

The numerical evaluation of the integral in Eq. (A10) was carried out by means of Simpson's rule. Tables of the functions θ and χ are available[34] for values of w up to 20 in steps of 0.5. For τ larger than about 2.0, these tables become unsatisfactory because $\sin w\tau$ and $\cos w\tau$ change so rapidly in this range that smaller increments are required. It is then necessary to compute the functions from Eqs. (15) and (18). Tables for the Bessel and Struve functions are available[35] for values of the argument up to 15 in steps of 0.01.

The numerical integration can be carried out only to some upper limit λ. The corresponding result of the calculation, including the constant term $\frac{1}{2}$ in Eq. (A10), is denoted by I_λ while the contribution to the integral from the range of integration λ to ∞ constitutes a remainder R_λ. An estimate for the latter may be obtained if the functions $J_1(w)$ and $\mathbb{S}_1(w)$ are approximated by the leading terms of their asymptotic representations.[15,33] The resulting expressions for $\theta(w)$ and $\chi(w)$ are given in Eqs. (19). It is then easily verified that the denominator of the integrand deviates from $4w$ by less than one-half percent for $w\geqslant10$. One obtains thus an estimate for the remainder

$$R_\lambda=-\frac{2}{\pi}\int_\lambda^{\infty}\frac{\cos w\tau}{w^2}dw+\frac{4}{\pi^3}\int_\lambda^{\infty}\frac{\sin w\tau}{w^3}dw. \qquad (A11)$$

By substituting a new variable for $w\tau$ and integrating by parts, one can reduce these integrals to sine integrals[36]

[34] See reference 11, p. 447.

[35] See, for example, reference 15, pp. 157 and 219.

[36] H. B. Dwight, *Tables of Integrals and Other Mathematical Data* (MacMillan Company, New York, 1947), revised edition, pp. 89

$\mathrm{Si}(\lambda\tau)$. Collecting all resulting terms yields

$$R_\lambda=\frac{\tau}{\pi}\left(1+\frac{\tau}{\pi}\right)\left[2\frac{\cos\lambda\tau}{\lambda\tau}+2\mathrm{Si}(\lambda\tau)-1\right]+\frac{2}{\lambda^2\pi^3}\sin\lambda\tau. \quad \text{(A12)}$$

In the numerical work, the value $\lambda=14.5$ was used and the results are shown in Table I. Although the value of R_λ is only an approximation, it is seen that even considerable errors would not affect the results significantly since this term represents only a minor correction of I.

APPENDIX III. CALCULATION OF THE ACCELERATION PRODUCED AT THE HEAD OF THE REFLECTED EXPANSION WAVE

The acceleration of a gas particle is given by the fundamental equation of motion as

$$\frac{Du}{Dt}=\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}=-\frac{1}{\rho}\frac{\partial p}{\partial x}=-\frac{2}{\gamma-1}\frac{\partial a}{\partial x} \quad \text{(A13)}$$

where, on the right-hand side, use has been made of the fact that all changes of state in an expansion wave are isentropic.

The pressure at the exit of the duct, following the arrival of the shock wave, is given by the data presented in Table I and plotted in Fig. 8. The region near the head of the reflected expansion wave is only affected by the initial portion of the pressure decay and, as seen from Fig. 8, this may well be represented by the linear relation

$$(p-p_0)/(p_1-p_0)=1-ka_0t_e/D. \quad \text{(A14)}$$

An excellent fit is obtained by setting $k=\frac{2}{3}$ [dotted line in Fig. 8(a)].

If Eq. (A14) is combined with the condition of isentropic changes of state in the expansion wave, one obtains

$$\frac{a}{a_1}=\left(\frac{p}{p_1}\right)^{(\gamma-1)/2\gamma}=\left\{\frac{p_0}{p_1}+\left(1-\frac{p_0}{p_1}\right)\times\left[1-\frac{ka_0}{D}\left(t_m-\frac{L}{a-u}\right)\right]\right\}^{(\gamma-1)/2\gamma}, \quad \text{(A15)}$$

where the lag time at the exit has been expressed by

$$t_e=t_m-L/(a-u) \quad \text{(A16)}$$

from Eqs. (3) and (7).

One can now take the logarithmic differential of Eq. (A15) but must remember that t_m is to be treated as a constant since the partial derivative with respect to x is required. Furthermore, one must set $dL=-dx$ because an increase of L corresponds to a decrease of the x coordinates which is measured positive in the direction of the flow (see Fig. 2).

The result of this differentiation is

$$\frac{da}{a}=-\frac{\gamma-1}{2\gamma}\frac{p_1}{p}\left(1-\frac{p_0}{p_1}\right)\frac{ka_0}{D(a-u)^2}\times[(a-u)dx+L(da-du)]. \quad \text{(A17)}$$

From the characteristic relation (4), it follows that

$$du=-2da/(\gamma-1)$$

and if this is substituted into Eq. (A17), one obtains after some rearranging

$$\frac{\partial a}{\partial x}=-\frac{(\gamma-1)ka_0a(a-u)(p_1-p_0)}{(\gamma+1)ka_0a(p_1-p_0)L+2\gamma(a-u)^2pD}. \quad \text{(A18)}$$

The conditions at the head of the expansion wave, are given by $a=a_1$, $u=u_1$, and $p=p_1$. If one also sets $k=\frac{2}{3}$ and combines Eqs. (A18) and (A13), one obtains the acceleration of a gas particle on entering the expansion wave as

$$\frac{Du}{Dt}=\frac{2a_0a_1^2(a_1-u_1)(p_1-p_0)}{(\gamma+1)a_0a_1(p_1-p_0)L+3\gamma(a_1-u_1)^2p_1D}. \quad \text{(A19)}$$

---

and 96. For tables of the sine integral see Mathematical Tables Project (Work Projects Administration, New York), "Tables of sine, cosine and exponential integrals" (1940) Vols. I and II, and "Tables of sine and cosine integrals for arguments from 10 to 100," (1942).

Domain Configurations and Crystallographic Orientation in Grain-Oriented Silicon Steel

W. S. PAXTON AND T. G. NILAN
United States Steel Corporation, Applied Research Laboratory, Pittsburgh, Pennsylvania
(Received January 21, 1955)

The magnetic domain structure of polycrystalline grain-oriented 3¼% silicon steel has been investigated by the magnetic-powder-pattern technique and the orientation of individual grains determined by the etch-pit optical-goniometer technique. A semiquantitative relationship has been found between the domain patterns and crystalline orientation. The observed powder patterns can be grouped into five types, depending upon the angle, ϕ, between the rolling plane and the <100> direction lying closest to the rolling plane. When the rolling direction is known and ϕ is $<10°$, semiquantitative grain-by-grain orientation information can be obtained solely from the domain patterns.

The behavior of two of the types of patterns in an applied magnetic field is shown, and on the basis of this behavior, a three-dimensional domain configuration is proposed for one of the less complicated surface patterns.

INTRODUCTION

UNTIL recently, investigations of the relationship between ferromagnetic domain configurations and crystallographic orientation in silicon steel or iron have been limited to the study of carefully grown and polished single crystals. The patterns on polycrystalline material have been considered too complex to provide information necessary for the formulation of a consistent theory of domain structure. In addition, the grain-by-grain orientation data required are difficult to obtain by standard x-ray techniques.

Some work on the domain structure of polycrystalline silicon iron has been reported in the literature. The domain structure of a polycrystalline sample under tension has been studied by Dijkstra and Martius.[1] The x-ray orientation information that they reported pertained only to the well-oriented grains examined. Bates and Hart[2] have reported a qualitative comparison between the domain patterns of a polycrystalline specimen and those of a single crystal. In their work, only the orientation of the single crystal was known.

Néel[3] has predicted from theoretical considerations the shape of domain patterns on the surface of iron crystals. The properties of silicon iron, such as magnetic anisotropy and saturation induction, are sufficiently like those of iron that the same patterns may be expected on the silicon iron. Williams, Bozorth, and Shockley[4] have studied and explained the patterns on single crystals of silicon iron. Following the methods of these workers, we have observed the domain patterns on polycrystalline silicon-steel samples, determined the orientation of the grains, and attempted to explain the surface patterns in terms of the domain configurations within the sample.

In addition to furthering the fundamental knowledge of the process of magnetization in ferromagnetic metals and alloys, such information can also be of great technological importance because elementary domain processes are, in the final analysis, responsible for permeability, hysteresis, and ac losses in magnetic materials.

EXPERIMENTAL WORK

The specimens used in this investigation were ½-inch disks, initially 0.014 inch thick, punched from commercially produced grain-oriented 3¼% silicon steel. This material was produced by the conventional method, which includes cold reduction to final thickness in two stages, each followed by an anneal in the temperature range 760° to 870°C, and finally a high-temperature anneal at about 1180°C.

The disks were mechanically polished through 3/0 metallographic paper, electropolished, and then lightly etched in a ferric sulfate solution to reveal the grain boundaries. By using a microscope with a micrometer stage, an accurate map of the grains in the specimen was drawn so that any given grain might be repeatedly located for obtaining both domain and orientation data on the same grains. A photomacrographic technique for grain mapping was tried at first but proved to be impracticable, since the deep etch required by this technique to clearly define the boundaries between grains of almost the same orientation made the disc unsuitable for further work.

After the area to be studied had been mapped, the specimen was repolished mechanically and electrolytically for observation of magnetic-powder patterns. Contrary to the findings of other workers,[1] we found it unnecessary to stress-relief anneal the specimens after repolishing because the patterns observed before and after such annealing were essentially the same.

Magnetic-powder patterns were obtained, with the technique of Williams, Bozorth, and Shockley,[4] by using the colloidal suspension of Fe_3O_4. The electrolyte used in electropolishing, a mixture of chromic and phosphoric acid, was the same as that used in their

[1] L. J. Dijkstra and U. M. Martius, Revs. Modern Phys. **25**, 146 (1953).
[2] L. F. Bates and A. Hart, Proc. Phys. Soc. London, **A66**, 813 (1953).
[3] L. Néel, Cahiers Phys. **25**, 1 (1944).
[4] Williams, Bozorth, and Shockley, Phys. Rev. **75**, 155 (1949).

work. We found that the electropolishing scum could be removed from the surface by cathodically cleaning the specimen in a 5% sodium hydroxide solution for a few seconds.

Patterns were observed under both dark- and bright-field illumination. Dark-field illumination, as shown by Williams, was best for patterns of simple geometry. For the more complicated patterns exhibited by grains with planes of high indices parallel to the sample surface, a judicious blending of both dark- and bright-illumination was used to reveal lines of symmetry not readily apparent under dark-field illumination alone. The photographs were made with a 35-mm camera on Kodak Microfile film. This film is very slow, but has the advantage of very fine grain and high contrast. Powder patterns of interest were photographed and their coordinates noted on the grain maps.

The orientation of the selected grains was determined by the etch-pit optical-goniometer method described by Barrett.[5] A commercial two-circle goniometer was used for these measurements. The specimen for the domain-configuration studies was again electropolished to remove traces of dried colloid and then etched in a solution containing 100 g of ferric sulfate and 10 ml of concentrated sulfuric acid per liter of water. The solution was kept at 80°C while in use.[6] The facets of the etch pits developed in silicon iron by this etchant are parallel to the {100} planes of the grain. An objectionable black deposit formed on the surface during etching, but this was removed by swabbing the sample with a 10% sodium nitrate solution containing a few drops of nitric acid per 200 ml of solution.

Examples of the etch pits are shown in Fig. 1. We cannot prescribe precisely an etching time that will invariably yield good results, inasmuch as the optimum

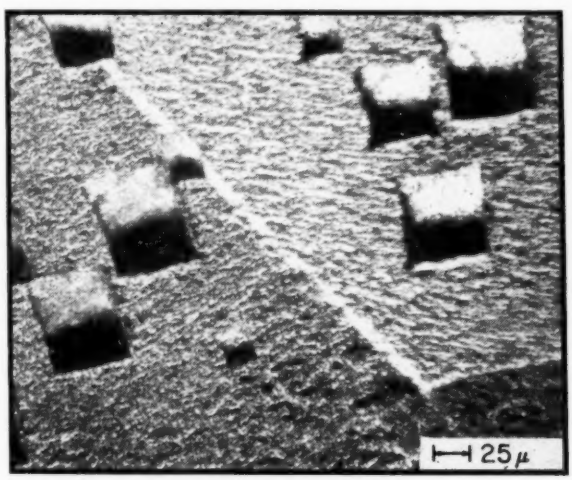


FIG. 1. Etch pits in two grains of silicon steel differing slightly in orientation.

[5] C. S. Barrett, *Structure of Metals* (McGraw-Hill Book Company, Inc., New York, 1943) pp. 173–177.
[6] P. Dunsmuir, Brit. J. Appl. Phys. **3**, 1 (1952).

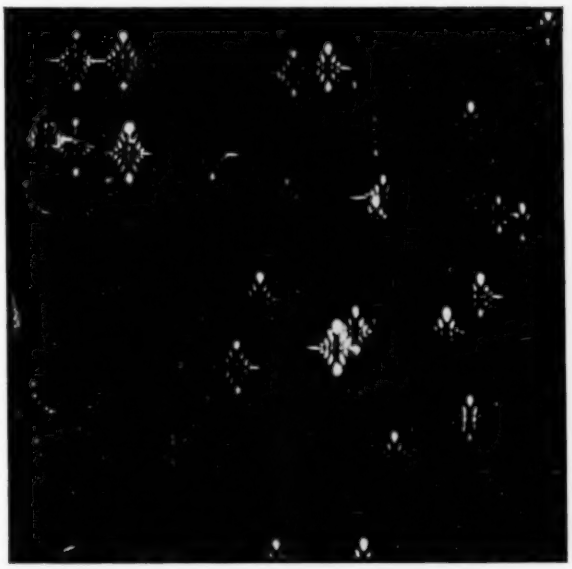


FIG. 2. Interference patterns formed by reflection from etch pits with the microscope of the optical goniometer slightly out of focus.

etching time varied from 1 to 10 minutes for different specimens from the same lot of material. It was necessary to remove the specimen from the etchant for examination at 1- to 2-min intervals to prevent overetching. With careful control, etch pits of a good rectangular shape could usually be obtained.

The goniometer is equipped with a telescope, focused at infinity, which can be converted to a microscope of approximately 15× magnification by placing a lens at the end of the telescope tube. The telescope arrangement is used, together with various apertures in the collimator, for determining the orientation of surfaces having high reflectivity. In most cases, the grains included in our study were too small in area to allow use of the telescope. Accordingly, we used the microscope arrangement, with the microscope focussed at the surface of the specimen. The intensity of reflection from the facets of the etch pits was found to have a broad maximum, which introduced an uncertainty of 2 to 3° in the angular readings. We found, however, that when the microscope was defocused slightly by moving the objective closer to the specimen, the points of light became interference patterns, as shown in Fig. 2. These patterns were symmetrical only when the grain was properly aligned. The orientations of several very large grains were determined by this method, and single-crystal pole figures were obtained for the same grains from x-ray diffraction back-reflection patterns. The agreement between the two sets of results was excellent.

By use of the interference patterns, we have accurately determined the orientation of poles that deviate from the normal to the specimen surface by less than 10°. We found that etchants other than ferric sulfate tended to etch the specimen so deeply that deter-

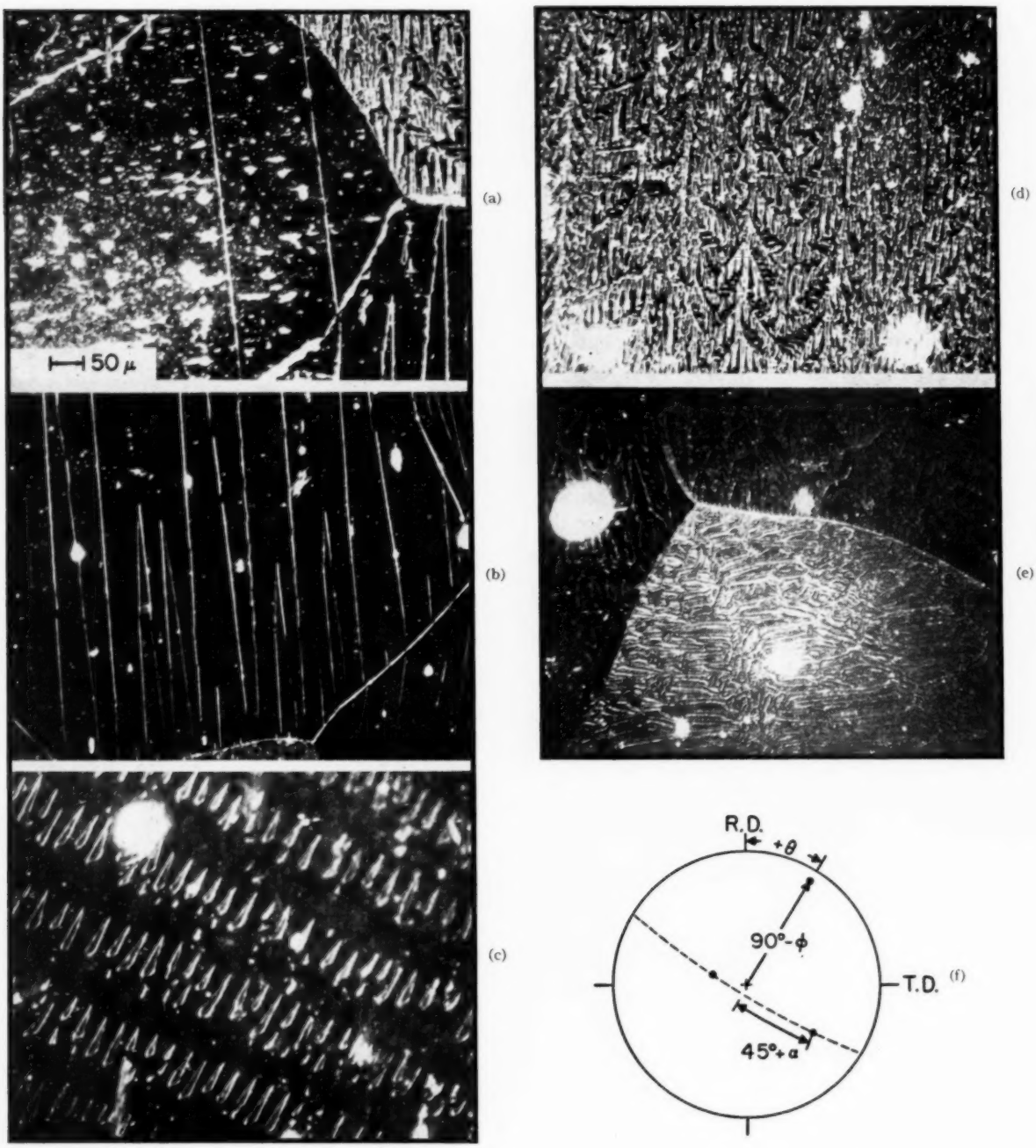


FIG. 3. Five basic types of domain patterns observed on demagnetized material. (a) Wall pattern $\phi=0°, \theta=-5°, \alpha=+3°$. (b) Spike pattern $\phi=+2°, \theta=-2.5°, \alpha=-14°$. (c) Lozenge pattern $\phi=+4°, \theta=+21°, \alpha=+10°$. (d) Lace pattern $\phi=+8.5°, \theta=+6°, \alpha=+3°$. (e) Maze pattern $\phi=+21°, \theta=-1°, \alpha=-5°$. (f) Typical pole figure for the {100} planes defining ϕ, θ, and α.

minations of poles within 15° of the normal was very difficult, because of the excessive background brightness of the specimen surface. By means of this technique, we have determined the orientation of grains as small as 5×10^{-2} mm² in area.

We have found it more convenient to plot the angular readings on a polar stereographic net rather than on the more commonly used Wulff net described by Barrett. With the polar net, it is not necessary to rotate the net. The pole is located on the net merely by plotting the horizontal-circle reading of the goniometer as the angular distance from the center of the net, along the radius corresponding to the vertical-circle reading.

CLASSIFICATION OF POWDER PATTERNS

Examination of magnetic-powder patterns on a large number of grains in various lots of silicon iron have indicated that the patterns on samples in the demag-

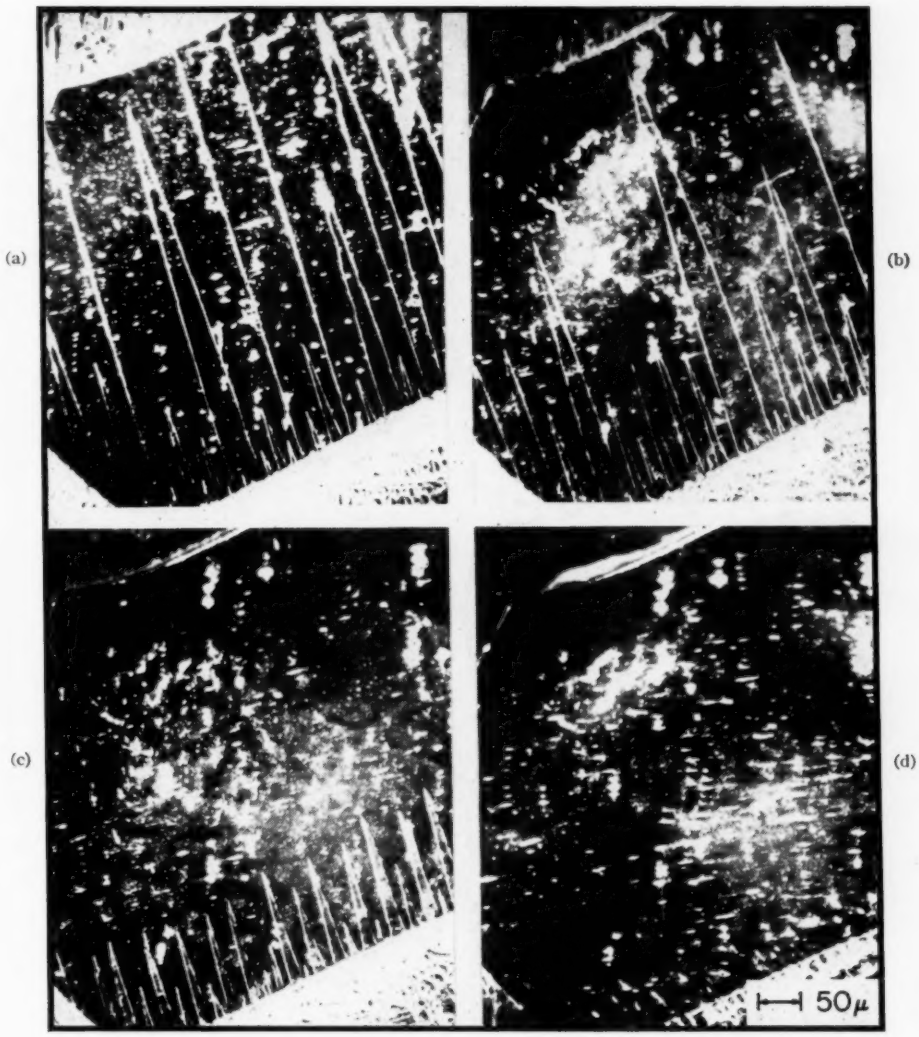


FIG. 4. Behavior of a spike pattern in a field applied parallel to the [100] direction. In (a) the grain is demagnetized. In (b) to (d) the field has been applied, increasing in size from (b) to (d). In (d) the grain has become saturated.

netized state could be classified into five types. For convenience in discussion, the descriptive terms "walls," "spikes," "lozenges," "lace," and "maze" have been applied to these five types of patterns. An example of each type is shown in Figs. 3(a)–3(e), together with a typical {100} pole figure showing the angular parameters used to describe the orientation of the grain. All these patterns have previously been described in the literature or have appeared in published photographs of domain structure. However, to our knowledge, no attempt has been made to classify the patterns by type, or to show a systematic relationship between the patterns and crystalline orientation.

It should be mentioned that our classification does not include the "tree" pattern described by Williams[4] for grains having a {100} plane close to the plane of the sample. This pattern is rarely found in the material of interest to us, which nominally has a (110) [100] texture. In addition, we are restricting this discussion to the dependence of domain structure upon only one ⟨100⟩ direction. The "tree" pattern would require consideration of the angular deviation of two ⟨100⟩ directions from the plane of the sample.

To clarify our classification further, we defined the types of patterns in the following manner:

"Walls" [Fig. 3(a)] are long, continuous, parallel boundaries usually widely spaced in comparison with the other patterns, and are often continuous across grain boundaries. The spacing between walls apparently varies with the grain size, the large grains having wider spacing between walls than the small grains. The observed spacings are generally in agreement with those observed by other investigators.

"Spikes" [Fig. 3(b)] are long, narrow, pointed figures originating at a grain boundary. In a grain having this pattern, all of the spikes point in the same

direction. In small grains, wall patterns of close spacing will often be found along with spikes. These walls will be discussed in greater detail later.

"Lozenges" [Fig. 3(c)] are small, pointed figures arranged in rows that are perpendicular to the direction in which the figures point. They are much smaller than spikes and are more numerous within a grain. These figures are very uniform in size, and the spacing within a row and between rows is very regular.

"Lace" [Fig. 3(d)] covers complicated patterns that still exhibit some over-all symmetry and regularity.

"Maze" [Fig. 3(e)] applied to patterns that do not exhibit any basic symmetry or regularity.

RELATIONSHIP OF DOMAIN TYPE TO CRYSTALLOGRAPHIC ORIENTATION

We have chosen three angular parameters to define the orientation of a given grain with respect to the customary reference plane and reference direction used in discussing sheet-steel product. The reference plane is called the rolling plane; the reference direction is called the rolling direction. The [100] direction of a grain is herein defined as that one of the ⟨100⟩ directions having the smallest angle of deviation from the rolling plane. We can define the angle θ as the angle formed by the rolling direction and the projection of the [100] direction on the rolling plane. A second angle, ϕ, is defined as the angle by which the [100] direction deviates from the rolling plane, that is, as the angle of tilt of the [100] out of the rolling plane. If a grain has a ⟨110⟩ direction in the surface, the [010] and [001] directions make equal angles with the rolling plane. The third angle, α, is defined as the angle of rotation of the grain about the [100] direction from the equiangular position.

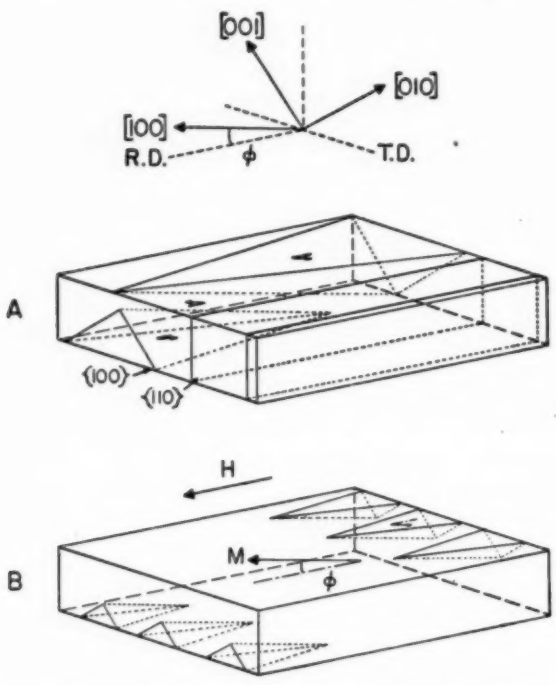


FIG. 5. Proposed models for three-dimensional configuration of domains represented by spike pattern. In (a) the grain is in the demagnetized state, corresponding to Fig. 4 (a), and some of the spikes have transformed into the narrow wall pattern. In (b), a field has been applied parallel to the [100] direction, corresponding to Fig. 4(c). Domain magnetizations are parallel to the [100] and [$\bar{1}$00] directions. Domain walls are parallel to the {100} and {110} planes as shown.

A comparison of orientation measurements with corresponding domain patterns revealed a remarkably good relationship between the type of pattern and the angle ϕ. It appeared that the angle α had little visible effect on the pattern, at least for small values of α. In the demagnetized material, the domain configuration appeared to be independent of the angle θ. In an investigation of over two hundred grains having orientations corresponding to wide ranges of both ϕ and α, no exception to this relationship was found.

On the basis of this investigation, we find that the type of domain pattern observed is determined by the magnitude of ϕ. When ϕ is in the range 0 to 2°, the grain exhibits a wall pattern [Fig. 3(a)]. The domain walls and the domain magnetization are parallel to the [100] direction. Under the influence of an applied field, the walls undergo translation in a direction perpendicular to the trace of the wall on the specimen surface.

A spike pattern is observed when ϕ is between 2 and 4°. The spikes point in the [100] direction, as shown in Fig. 3(b). When changes in this pattern are observed, as the specimen is taken from the demagnetized state to saturation and through a hysteresis loop with the applied field (H) parallel or nearly parallel to [100], it is apparent that the spikes are outlines of a basic structure [Fig. 4(a)–4(d)]: that is, they appear to be domains of reverse magnetization.

In recent publication, Goodenough[7] has calculated the nucleation energies for domains of reverse magnetization at grain boundaries and at crystal surfaces. He concludes that the probability for creation of such nuclei is largest at grain boundaries and depends on the relative angles of orientation of the axes of magnetic symmetry of the contiguous grains. Our results are evidently not in accord with the findings of Goodenough since our work indicates that the domains of reverse magnetization are characteristic of the absolute orientation of the grain and are not influenced by the orientation of the neighboring grain, at least in the demagnetized state.

The behavior of the spike pattern under an external magnetic field is as follows:

When the strength of the applied field is decreased from saturation, small spikes form at one grain boundary, as shown in Fig. 4(c). During a further decrease of field strength, these spikes grow until they suddenly change to walls that are parallel to the [100] direction, as shown in Fig. 4(b).

If the applied field is reversed and increased, after

[7] J. B. Goodenough, Phys. Rev. **95**, 917 (1954).

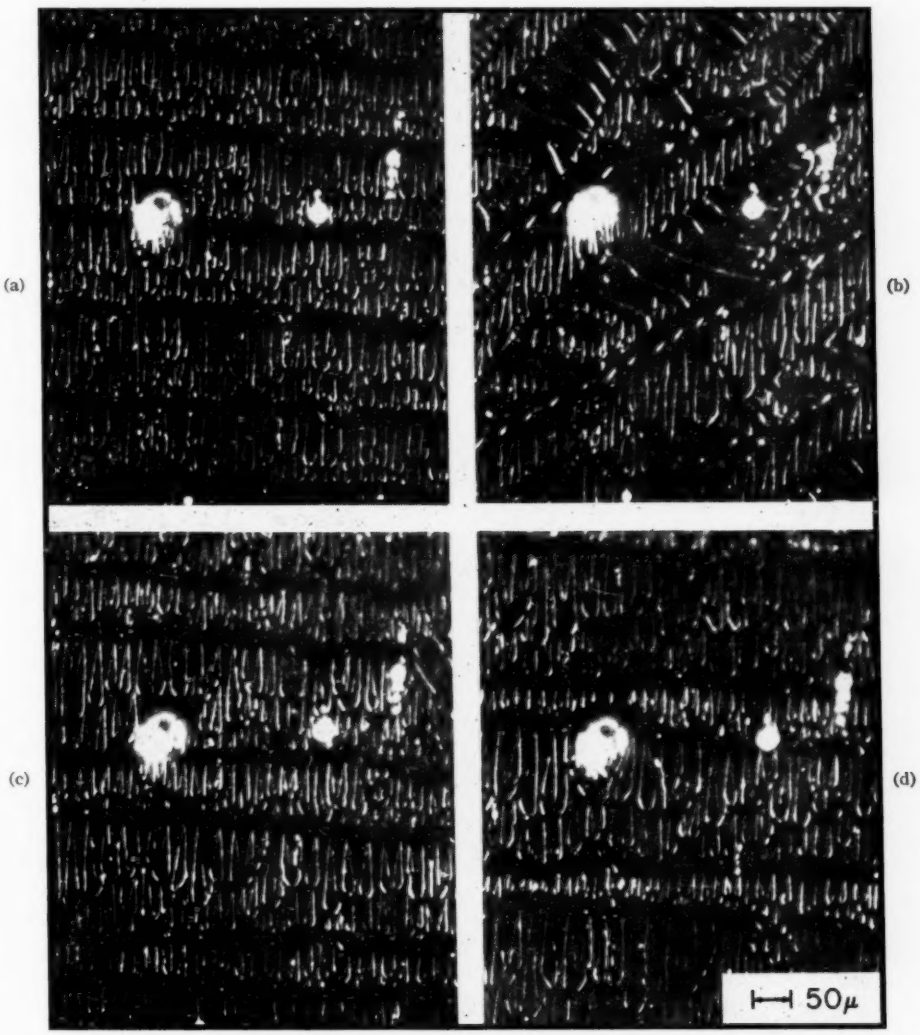


FIG. 6. Effect of a field applied parallel to the [100] direction on the lozenge pattern. In (a) the sample is demagnetized. In (b) a field has been applied and the grain is approaching saturation. In (c) the applied field is smaller than that in (b). In (d) the field in (c) has been reversed. The etch pits in the surface of the grain serve as reference marks.

the spike-to-wall transition has occurred, the walls return to a spike configuration. However, one difference in pattern is noticed. Prior to the spike-to-wall transition, a given boundary, forming one side of a spike, makes an acute angle with the grain boundary, which serves as the base line. After the second transition (wall-to-spike), this boundary makes an obtuse angle with the same base line. This re-formed spike decreases in size and finally disappears as saturation is reached.

The grain size of the material investigated was sufficiently large, relative to the sheet thickness, that the grain boundaries penetrated the sheet almost perpendicularly and thus preserved the grain shape from side to side almost unchanged. Consequently, it was possible to observe the patterns on the opposite surfaces of the same grains. In this way, it was learned that the spikes on one surface of a grain pointed in a direction 180° to those on the opposite surface. On one surface of the specimen, the spikes originated at, and pointed away from, a given grain boundary, whereas on the opposite surface of the specimen, the spikes pointed toward this boundary and originated on the boundary across the grain.

In the models proposed to explain these configurations (Fig. 5), we have assumed that the walls of the domains, as shown in Fig. 4(c), were parallel to (010) and (001) planes. There is justification for this assumption, since the calculation of ϕ based upon the domain patterns is in good agreement with the measured value. The calculation is made in the following manner. If the area outlined by the spike is treated as one face of a tetrahedron, with {100} planes as the two other faces, a cross section of the domain transverse to the rolling direction is an isosceles triangle. Since the (001) and

(010) planes make an angle of 45° with the surface of the grain, it is evident that the altitude of the triangle is equal to one-half the width of the spike at the point where the cross section was taken. If a cross section of the domain parallel to the rolling direction is next considered, this section constitutes a right triangle, with ϕ being the included angle between the hypotenuse and that side of the triangle representing the surface of the grain. The other side of the triangle is the altitude of the triangle formed by the transverse cross section, and it is therefore equal to one-half the width of the spike at the grain boundary. Angle ϕ is then the angle the tangent of which is the half-width of the spike divided by the length of the spike.

Measurements on the spikes in Fig. 3(b) gave a value of 2°30′ for ϕ. The value of ϕ determined by the optical goniometer for this grain was 2°.

The lozenge pattern, shown in Fig. 3(c), is observed when ϕ is in the range 4 to 7°. Two dominant lines of symmetry are apparent. The lozenge figures, which resemble small spikes that are either open or closed at the base, point in the [100] direction. In addition, the figures are arranged in rows parallel to the $[01\bar{1}]$ direction. Both the spacing between lozenges in a row, and the spacing between rows, are regular.

The behavior of this configuration under an applied field (H is approximately parallel to the [100] direction) is shown in Figs. 6(a)–6(d). In the demagnetized state [Fig. 6(a)], the lozenges are all of a uniform size. As the strength of the applied field is increased, figures in alternate rows either increase or diminish in size. When the field strength is increased further, the pattern changes to the "echelon" array of lozenges shown in Fig. 6(b). This pattern in turn changes to the pattern Bates has called "tadpoles"[2] when the field strength is again increased. In Figs. 6(c) and 6(d) the sample is shown at some level of induction intermediate to Figs. 6(a) and 6(b). The field applied in Fig. 6(c) has been reversed in Fig. 6(d). The rows of figures that were large in Fig. 6(c) are small in Fig. 6(d) and vice versa. The etch pits serve as reference marks to identify the rows of lozenges.

Isolated lozenge figures have been seen in the vicinity of voids or surface irregularities on grains in which the dominant pattern was spike or wall. In these instances, the lozenges still point in the [100] direction. From this observation, we judge that the lozenge figures are a flux-closure pattern.

When ϕ is in the range 7 to 13°, the lace pattern appears. The maze pattern is observed when ϕ is greater than 13°. No attempt has been made to develop models to explain these more-complex configurations.

We have been able to use the domain configurations on a polycrystalline sample of grain-oriented silicon steel to obtain semiquantitative information on the grain-by-grain orientation of the specimen. The value of ϕ can be estimated closely from the type of domain pattern observed, at least, for the wall, spike, and lozenge configurations. Even for the lace pattern, a fairly good estimate of the value of ϕ can be made, since the patterns increase in complexity as ϕ increases. If the rolling direction is known, θ can be determined by measuring the angle between the rolling direction and the [100] direction. In the spike and lozenge patterns, the figures point in the [100] direction, so that the measurement of θ is straightforward. It is sometimes possible to determine θ for a grain showing the lace pattern, if ϕ is less than 10°. In this case, the lace pattern usually contains some small, pointed figures, as shown in Fig. 3(d), that point in the [100] direction. Plots of the distribution of ϕ and θ as determined from the domain patterns showed excellent agreement with plots of the distribution of ϕ and θ as obtained from optical-goniometer measurements of etch-pit orientations.

ACKNOWLEDGMENTS

We wish to acknowledge many helpful discussions with G. E. Pellissier, in charge of the Physical and Analytical Methods Division of the United States Steel Corporation Applied Research Laboratory, and with Dr. J. E. Goldman, of Carnegie Institute of Technology. The x-ray orientation measurements were made by B. G. Reisdorf and E. E. Wicker.

Maximum Superheating of Water as a Measure of Negative Pressure

LYMAN J. BRIGGS
National Bureau of Standards, Washington, D. C.
(Received January 12, 1955)

Water superheated to the limit in an open tube develops an *internal* negative pressure P_n (tending to tear the molecules apart) equal to the saturation vapor at that temperature, less one atmosphere. Kendrick, Gilbert, and Wisner heated water in an open thin-walled capillary U-tube to 270°C for 5 sec before it exploded ($P_n = 53$ atmos). Using their method, the writer heated three tubes to 264°, 266°, and 267°C for 5 sec or more before explosion occurred ($P_n = 48$ to 51 atmos). The cohesive strength of water is thus sufficient to withstand an *internal* negative pressure of over 51 atmos at 267°C. At this temperature an additional negative pressure (applied *externally* through centrifugal force) would rupture the water column. It has been predicted from van der Waals' equation that the *external* negative pressure which the system could withstand would vanish at 273°C, in fair agreement with experiment.

INTRODUCTION

IT has long been known that if the dissolved gas in water is greatly reduced and no sharp points or other "boiling promoters" are present in the container, then the water can be superheated far above its normal boiling point before a bubble is formed and the water violently explodes. It occurred to me that superheating might be used as a measure of the negative pressure which water can withstand at high temperatures. Water in a sealed superheated tube at T° is prevented from boiling by the pressure of the saturated vapor above it. If the tube is opened and the water does not explode, it must now be under an *internal* negative pressure P_n (tending to tear the molecules apart) numerically equal to the vapor pressure at T° (in atmos) less 1 atmosphere. When the temperature is raised to a point where the forces of cohesion no longer equal or exceed the increasing negative pressure, the system explodes.

I found no mention of this concept in the various papers on negative pressure which have appeared since the turn of the century. But a brief comment by Osborne Reynolds[1] led me to Clerk Maxwell's *Theory of Heat*, 1871, and there I found that Maxwell had the same concept. He says: "In an experiment due to Defour, the water, instead of being allowed to touch the sides of the vessel, is dropped into a mixture of linseed oil and oil of cloves, which has nearly the same density as itself. By this means drops of liquid may sometimes be observed swimming in the mixture at a temperature of 356°F (180°C). The pressure of aqueous vapor at this temperature is nearly 10 atmos or about 147 lb weight on the square inch. Hence the cohesion of the water must be able to support at least 132 lb weight on the square inch (9 atmos)."

It should be noted that the *internal negative pressure* in a superheated liquid is the quantity that is measured. At the highest stable temperature it provides a measure of the cohesive strength of the liquid under the prevailing conditions.

Superheating thus provides a way to measure the limiting negative pressure of water at a temperature far above 100°C. Previous measurements do no appear to have been made at temperatures above 73°C by the Berthelot method[2] and above 50°C by the centrifugal method.[3]

EARLIER MEASUREMENT OF THE SUPERHEATING OF WATER

E. N. Harvey, A. H. Whiteley, W. D. McElroy, D. C. Pease, and D. K. Barnes[4] found that water could be superheated to 202°C–206°C at atmospheric pressure provided it was first subjected to a hydrostatic pressure of 1000 atmos or more for 15–30 minutes, in order to force the gas nuclei into solution.

F. B. Kenrick, C. S. Gilbert, and K. L. Wismer[5] measured the maximum superheating of a number of liquids, including water, in open capillary U-tubes. If the liquid did not explode after the lower part of the U-tube had been immersed in the bath for 5 sec, the tube was withdrawn, the bath temperature raised, and the process repeated. They obtained the highest value so far reported for the superheating of water, 270°C at atmospheric pressure. They do not state whether their thermometers had been standardized.

NEW MEASUREMENTS

Before repeating the superheating experiments of Kenrick *et al.* on water, I wished to confirm their conclusion that the liquid under test actually reached the temperature of the bath in an immersion time of only 5 seconds. To determine this, a butt-welded chromel-alumel thermocouple of fine wire[6] (0.14 mm diam) was threaded through an open capillary U-tube with dimensions approximating the U-tubes used by Kenrick *et al.*, and later by myself.

Silicone oil sufficient to cover the junction was put into the U-tube, which was then placed in a stirred bath

[1] Osborne Reynolds, *Collected Papers* (Cambridge University Press, London, 1900), Vol. 1, p. 241.

[2] H. H. Dixon, Proc. Roy. Soc. (Dublin) **12**, 60 (1901).
[3] L. J. Briggs, J. Appl. Phys. **21**, 721 (1950).
[4] Harvey, Whitely, McElroy, Pease, and Barnes, J. Cellular Comp. Physiol. **24**, 25 (1944).
[5] Kenrick, Gilbert, and Wismer, J. Phys. Chem. **28**, 1297 (1924).
[6] Constructed by the Pyrometry Laboratory, Natl. Bur. Standards.

at 240°C, where the emf of the thermocouple was balanced by a potentiometer. When the U-tube was removed from the bath, allowed to cool and again inserted, the potentiometer balance was always restored (to within 0.5°C) in 3 or 4 seconds. An immersion time of 5 sec is, therfore, sufficient to heat the water in the U-tube up to the bath temperature, but it is subjected to this maximum temperature for only 1 or 2 seconds.

The bath container consisted of a glass test tube (3 cm bore, 15 cm long) with a nichrome heating element wound spirally outside. It was thermally insulated sides and bottom, the temperature being controlled by a Variac. The bath liquid was a silicone high-vacuum diffusion-pump fluid,[7] which proved to be very suitable for a high-temperature liquid bath. It has a boiling point of 450°C, is clear, colorless, of low viscosity, apparently nontoxic, does not break down on heating, and evaporates very slowly even at 270°C. The bath was stirred by a propeller driven by a little 600 rpm synchronous motor, with a shaft 20 cm long to provide ample clearance. The mercurial thermometer used to measure the bath temperature had an immersion point of 5 cm, and was graduated in 2-degree intervals from 100° to 370°C. It had been standardized in the NBS thermometer laboratory. The immersion time was determined by counting in unison with one-second signals from a clock.

The capillary U-tubes were drawn from clean glass tubes (Pyrex brand), the stock tube (10 mm) being left attached at right angles to one end of the U-tube to serve as a handle. A one-cm length of the free end of the capillary was bent at right angles to the vertical U-tube to facilitate filling. The tubes were drawn just before using and were partially filled by capillarity with water that had been freshly boiled to about three-fourths of its original volume.

Over 30 tubes were drawn and tested, the bore diameter ranging from 0.2 to 0.5 mm. Three tubes gave the highest values, I observed, for the superheating of water for 5 sec before explosion, namely 264°, 266°, and 267°C. The best tube had a bore diameter of 0.28 mm; wall thickness 0.22 mm. It withstood successive 10 sec immersions at about 4° intervals as the bath was raised from 211° to 267°, when it exploded after 7 sec immersion; refilled, 3 sec at 268°; refilled again, 8 sec at 266°C.

The other U-tubes showed a wide dispersion, ranging from 245° to 200°C. In view of the high values obtained for the three tubes, the failure of the others is probably to be ascribed to a surface defect on the wall of the capillary. D. A. Glaser[8] has shown that cosmic rays will liberate bubbles in a superheated liquid. However, the capillary U-tubes are so small that the probability of one being struck by an ionizing cosmic particle in 5 sec is of the order of only 1 part in 1000. But it is known that redistilled water still contains traces of ionizing particles, as does a glass container, and the measurements must thus be considered as representing a lower boundary.

Freshly-drawn capillaries gave the best results and often (but not always) the first filling withstood a higher temperature than later refills. Mote-free water[9] and water boiled under reduced pressure and then cavitated until no bubbles appeared did not prove to be better than boiled water, but only a few tests were made. A platinum capillary gave lower values than Pyrex brand glass.

The following change in technique was tried. A freshly drawn capillary was filled with boiled water and six sections were cut from the tube, each about 1 cm long. These were put into the stirred silicone oil bath and slowly heated. One section exploded at 190°C; a second at 193°; a third at 202°C, which set off a violent simultaneous explosion of all the other sections. Here the immersion time was much longer than that used in the U-tube measurements.

Kenrick *et al.*[5] have shown that when ether is superheated nearly to the maximum temperature it can withstand, the immersion time required for explosion on successive trials in the same tube at the same temperature varies greatly, ranging from 2 to 60 seconds. It seems not unreasonable to expect such a chance variation when we consider the great increase in the momentum of the molecules at this higher temperature. If three or more of these molecules happen to collide at a point and then rebound elastically, it is quite conceivable that they would open up a void that would lead to the explosion of the system. The higher the temperature, the greater is the probability that such an event would occur in a specified time.

The results of the various measurements of the superheating of water, as already described, are summarized in Table I.

In the fourth column of the table is given the corresponding saturation vapor pressure in atmospheres for each temperature, rounded off from the accurate measurements by N. S. Osborne, H. F. Stimson, and D. C. Ginnings.[10] In accordance with the concept

TABLE I. Superheating of water in a container open to the atmosphere.

Observer	Method used	Temperature attained	Sat. vapor pressure atmos	Negative pressure atmos
Defour	Droplets in oil	180°C	9.9	9
Harvey *et al.*	Gas nuclei removed by hydrostatic pressure (1000 atmos)	206	17.4	16
Kenrick *et al.*	Capillary U-tube	270	54.2	53
Briggs	Capillary U-tube	264	49.2	48
		266	50.7	50
		267	51.8	51

[7] Dow Corning 703 fluid.

[8] D. A. Glaser, Phys. Rev. **87**, 655 (1952); **91**, 762 (1953).

[9] Prepared by the Pure Substances Laboratory, Natl. Bur. Standards.

[10] Osborne, Stimson, and Ginnings, J. Research Natl. Bur. Standards **23**, 261 (1939).

already presented, this vapor pressure, reduced by one atmos, gives the negative pressure (column 5) which the water was sustaining at the corresponding temperature. Strictly speaking, this internal negative pressure arising from molecular momentum should be *added* as a correction term to that measured directly by centrifugal force. For temperatures below 100°C, the correction for water is less than one atmos and may be neglected. For higher temperatures, it begins to enter, but no direct measurements exist to which to apply it. At 267°C it becomes the dominant factor, and we can say that water is capable of withstanding a negative pressure of *at least* 50 atmos at 267°C.

COMPARISON WITH MEASUREMENTS AT LOWER TEMPERATURES

A comparison of the limiting negative pressure of water near 270°C with that observed at lower temperatures is shown graphically in Fig. 1. The graph from

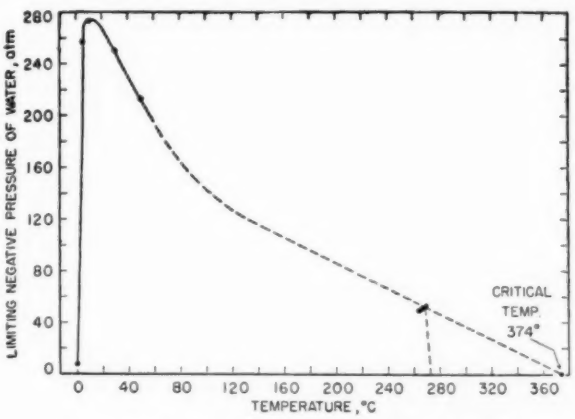


FIG. 1. Limiting negative pressure of water as a function of temperature.

0° to 50° is taken from the writer's measurements in which centrifugal force was used to stress the liquid.[3] The maximum negative pressure observed was about 270 atmos at 8°C. As the temperature was lowered, the negative pressure dropped precipitously, until at 0°C it was only about 5% of its maximum value. The measurements on the superheating of water (Table I) provide additional points on the graph at very high temperatures (264°–270°C, 48–53 atmos). The graph has been arbitrarily extended by a broken straight line to zero at the critical temperature (374°C), as a guide in establishing the slope of the extrapolated negative pressure curve as it approaches 270°C. The system abruptly vanishes at some point above 270°C, as indicated by the second broken line to zero pressure at 273°C. Between 100° and 270°C it should be capable of standing *some* tension (which might be applied centrifugally) even though it is already superheated, but this prediction has not been experimentally checked.

TABLE II. Highest temperature at which a liquid can sustain an external negative pressure. Comparison of theory and experiment.

Liquid	Crit. temp.	Ext. neg. pressure vanishes Computed	Observed
Ethyl alcohol	243°C	162°C	201°C
Benzene	289	201	203
Chloroform	263	179	173
Ethyl ether	194	121	143
Carbon disulphide	273	188	168
Sulphur dioxide	157	90	50
Methyl alcohol	240	160	180
Water	374	273	270

ZERO NEGATIVE PRESSURE AS DERIVED FROM VAN DER WAALS' EQUATION

J. Larmor[11] and later H. N. V. Temperley[12] have shown that if the van der Waals equation continuously applies, the *external* negative pressure which a liquid can withstand will vanish when the temperature equals or exceeds 27/32 of its absolute critical temperature; which, for water is 273°C.

Let us consider an idealized apparatus, consisting of a perfect cylinder, closed at one end and fitted with a frictionless, leak-proof piston. Let the enclosed space be completely filled with pure gas-free water, which thoroughly wets the walls of the cylinder and the face of the piston. Now pull outward on the piston. The water is subjected to a negative pressure. Van der Waals' equation predicts that at 273°C the *external* negative pressure required to disrupt the system has decreased to zero; at any higher temperature the water explodes into vapor.

This prediction is in fair agreement with the highest temperatures observed in the superheating measurements, namely 264, 266, 267, and 270°C. But this does not mean that the negative pressure of the system as a whole approaches the vanishing point asymptotically. Only the applied *external* negative pressure required to disrupt the system behaves in this way. At 270°C the negative pressure is 53 atmospheres. At a slightly higher temperature the system explodes. The negative pressure-temperature graph terminates abruptly. The whole system vanishes.

The remarkable deduction from van der Waals' equation that the external negative pressure a liquid can sustain vanishes when it reaches a temperature equal to 27/32 of its absolute critical temperature applies to any liquid. While the agreement with experiment is satisfactory for water, there is a marked discrepancy in the case of some other liquids, as is shown in Table II, in which the superheat measurements of Kenrick *et al.*[5] are used to provide the experimental data. (See also Temperley[11].)

I am indebted to Professor H. N. V. Temperley and Dr. N. E. Dorsey for helpful discussions of this subject.

[11] J. Larmor Proc. Lond. Math. Soc. (2) **15**, 191 (1916).
[12] H. N. V. Temperley, Proc. Phys. Soc. (London) **49**, 203 (1947).

Photoelectric Work Functions of the Borides of Lanthanum, Praseodymium, and Neodymium

RICHARD W. DECKER* AND DEAN W. STEBBINS
Iowa State College, Ames, Iowa
(Received February 14, 1955)

The DuBridge variation of the Fowler theory of photoelectric emission was utilized to determine the work function of the borides of La, Pr, and Nd. The work functions of LaB_6, PrB_6, and NdB_6 were found to be 2.74 ev, 3.12 ev, and 4.57 ev, respectively.

INTRODUCTION

MANY types of substances have been used to meet the demand for high thermionic emission current densities. In 1950, Lafferty[1] began a study of the emission properties of some rare earth boride cathodes. In his work he attempted to find an efficient electron emitting surface with a reasonable life. He obtained the photoelectric work function of lanthanum boride to be 2.67 ev by the analysis of the Fowler plot. Additional rare earth borides were examined here. Since some rare earths are available only in small quantities, and in Lafferty's method large amounts of the rare earth boride are required, a different procedure was devised for the present work, such that measurements could be made of the emission properties of these rare earth borides.

Samples of lanthanum boride were obtained from the Ames Laboratories of the U. S. Atomic Energy Commission and from the General Electric Laboratories. X-ray analysis indicated the crystal structure of the samples from the two sources was the same. The work function was obtained by the analysis of the DuBridge variation of the Fowler plot. DuBridge[2] examined the Fowler theory and suggested a method similar to Fowler's for obtaining the true photoelectric threshold. It has the advantage of employing a single frequency of incident light. DuBridge applied this method to data taken by several experimenters[3–5] and concluded the results obtained by each method were essentially the same. Because of the obvious experimental advantage of the DuBridge method, it was used in this work to obtain the work function of LaB_6, PrB_6, and NdB_6.

EXPERIMENTAL PROCEDURE

The work function of some rare earth borides was measured by new techniques which require only small amounts of the substance under investigation. The work function of LaB_6 obtained by this method was compared with that obtained by Lafferty to verify this technique.

The surface, on which these measurements were made, consisted of a tantalum ribbon 2 mm×30 mm ×0.025 mm, coated with the rare earth boride. The filament was first sandblasted with No. 120 carborundum and then carburized by heating to a bright red for a few seconds in an atmosphere of illuminating gas. The carburizing prevented diffusion of boron into the tantalum. A paste of finely ground rare earth boride and distilled water was then spread on the filament. An additional drop of water was added, and then the filament was heated in a small flame. The convection currents set up in the drop of water smoothed the boride surface as it dried.

The experimental tube, Fig. 1, was constructed of Corning Pyrex 7052 glass with the exception of a one-inch circular quartz window which was attached with a graded seal. The filament was supported by two 120 mil tungsten leads which also carried the filament heating current. Concentric glass tubing was affixed to each lead to increase the resistance between the plate

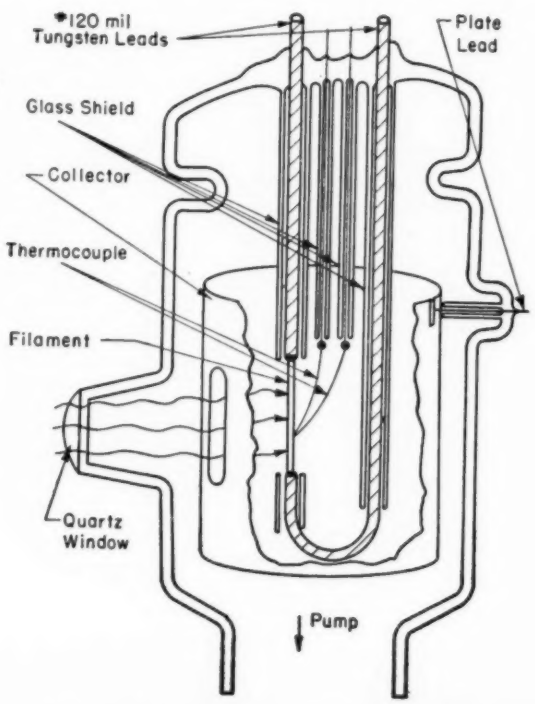


FIG. 1. Cut-away diagram of head assembly.

* Now at the Westinghouse Research Laboratories, East Pittsburgh, Pennsylvania.

[1] J. M. Lafferty, J. Appl. Phys. **22**, 299 (1951).
[2] L. A. DuBridge, Phys. Rev. **39**, 108 (1932).
[3] R. P. Winch, Phys. Rev. **27**, 1269 (1931).
[4] L. W. Morris, Phys. Rev. **37**, 1263 (1931).
[5] L. A. DuBridge and W. W. Roehr, Phys. Rev. **39**, 99 (1932).

and the filament, and to minimize the photoemission from the tungsten leads.

The temperature of the filament was measured by a chromel-alumel thermocouple held against the filament by the spring action of its leads. Monochromatic radiation was obtained from a Gaertner quartz monochromator using radiation from a Hanovia quartz mercury arc. The power for the filament and the mercury arc was supplied by a Sorrenson voltage regulator. The intensity of the source was held constant within 1%. The photoelectric current was measured by an electrometer, especially constructed for the Ames Laboratory of the U. S. Atomic Energy Commission. After a warmup period of two days the electrometer was found to drift less than 1% in 12 hours. Input resistors ranging in factors of 10 from 10^8 to 10^{14} ohms provided sensitivities of 10^{-8} to 10^{-14} amp. The output meter was calibrated and found to be linear within 2% on all ranges.

The prepared filament was mounted on a heavy tungsten lead as shown in Fig. 1, and the assembly was then fused to the vacuum system and evacuated. The vacuum system was all glass with a liquid nitrogen trap and a mercury diffusion pump backed by a mechanical fore pump. During the first day of evacuation, the vacuum system was heated to approximately 420°C. The filament was kept at a temperature of approximately 650°C. This higher temperature aided degassing the filament and also hindered contamination. The baking period lasted 26 hours for each run. The filament was kept at an elevated temperature throughout the experiment. This method of evacuation yielded pressures of less than 10^{-7} mm of Hg. Measurements were begun on the second day of evacuation. The work function was found to decrease for a period of two to three days and then to remain at the reported value. Values obtained for the work functions of LaB_6, PrB_6, and NdB_6 for initial runs after the baking period, were 3.5, 3.7, and 5.2 electron volts, respectively.

A definite procedure was devised and followed in order to obtain the cleanest surface possible. The assembly was first degassed as noted in the foregoing, and the filament then flashed for one minute at 1400°C before each measurement. The measurement of the emission was taken when the filament returned to an equilibrium temperature T, as indicated by the thermocouple reading. This was repeated for six to ten different temperatures for each frequency of incident light.

The number of measurements taken in each run depended upon the temperature range that could be used. The temperature range was limited by the highest temperature at which the filament could be raised and still distinguish the photoelectric current from the thermionic current. Since the thermionic emission increases logarithmically with the temperature, it soon becomes large compared to the photoemission. At the highest temperature used, the photoemission was about 25% of the total emission. Depending on the magnitude

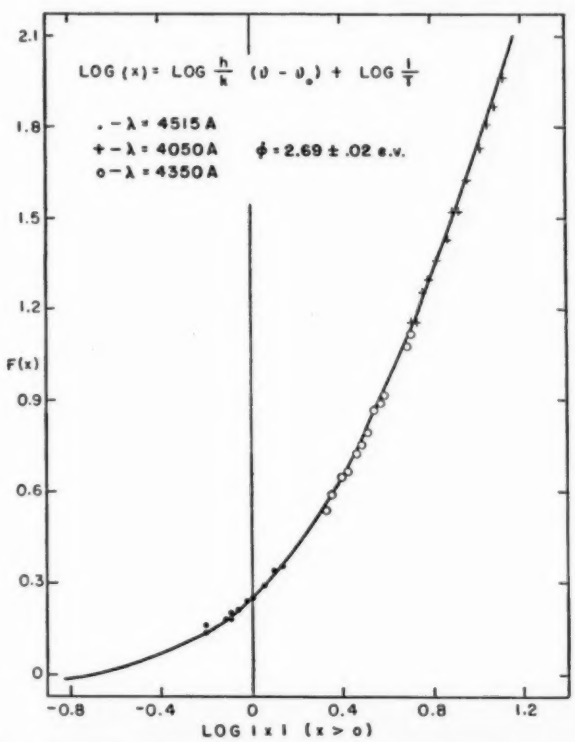


FIG. 2. Comparison of the theoretical curve with the observations for LaB_6 from General Electric Laboratories.

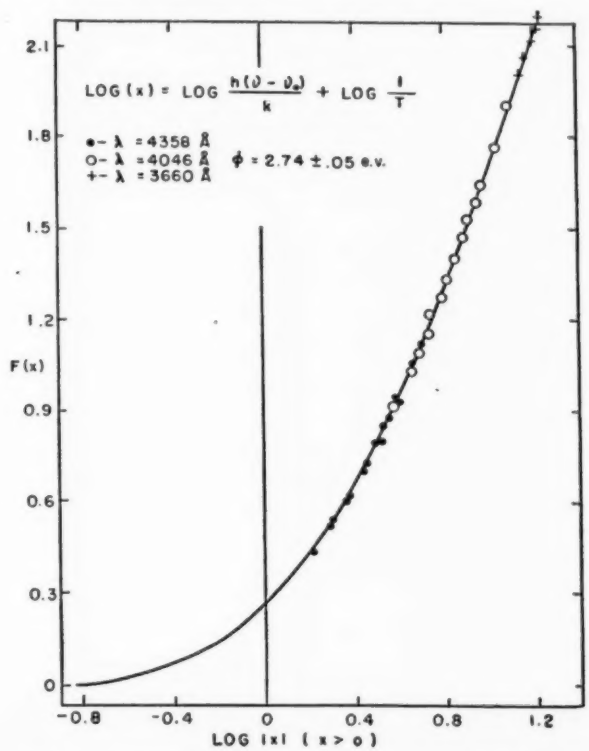


FIG. 3. Comparison of the theoretical curve with the observations for LaB_6 made at Iowa State College.

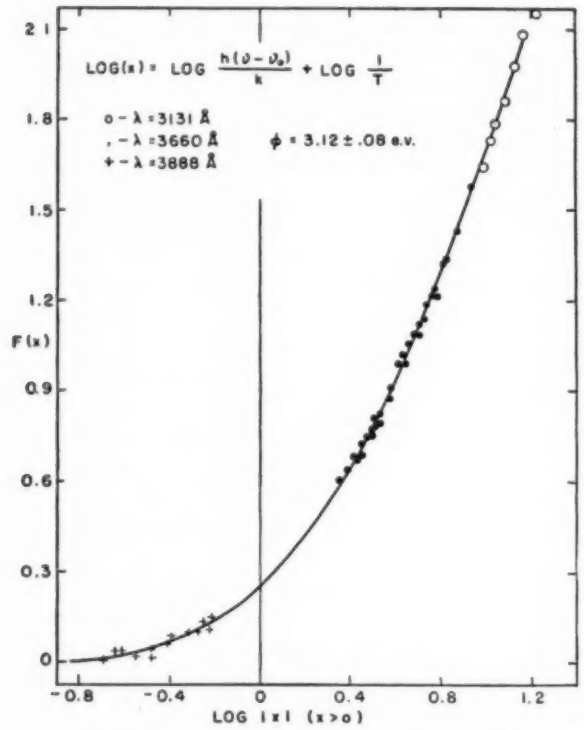


FIG. 4. Comparison of the theoretical curve with the observations for PrB_6.

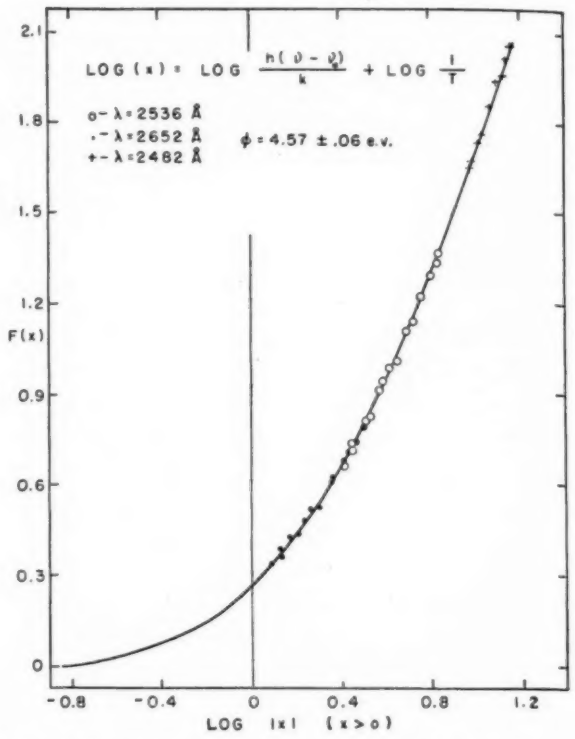


FIG. 5. Comparison of the theoretical curve with the observations for NdB_6.

of the photoemission, the temperatures that were used ranged from 298°K to 980°K.

The temperature and emission data for each frequency of incident light were plotted as log I/T^2 *versus* log $1/T$. The shift necessary to superimpose this curve on the theoretical curve[6] was then determined. Finally, the theoretical curve and the data (after applying the shift) were plotted on the same axes. The final plots are shown in Figs. 2, 3, 4, and 5.

RESULTS

The work function obtained for LaB_6 sample obtained from the General Electric Laboratories is 2.69±0.02 ev. This value compares very well with 2.67 ev obtained by Lafferty.

The work function of the sample of LaB_6 produced at Iowa State College was 2.74±.05 ev. It is thus seen that the value of 2.74 ev, 2.69 ev, and 2.67 ev are within the combined probable error. The agreement of the three values indicates that the sample produced at Iowa State College is probably the same as the General Electric sample.

The work function of PrB_6 was found to be 3.12 ±0.08 ev.

The work function of NaB_6 was found to be 4.57 ±0.06 ev.

The average probable error for work functions measured in this investigation is about 3%.

[6] L. A. DuBridge, Phys. Rev. **39**, 108 (1932).

DISCUSSION

In the work of Lafferty, the work function was measured using a solid bar of LaB_6 180 mils thick. This can then be assumed to be an infinitely deep substance, and the work function obtained can be assumed to be the work function of the material.

In the work described here, the work function was measured from a thin coating on a carburized tantalum ribbon. The results of Lafferty and the results of the authors for the work function of LaB_6 are in agreement. Therefore, it is concluded that the base metal of the filament has no effect on the work function for coatings of this thickness. The thickness of the coatings in this investigation was about one mil.

SUMMARY

The work functions of the two rare earth borides were measured by a photoelectric technique. The technique was verified by a comparison with measurements from a known sample. Useful techniques learned in preparing coated cathodes were described.

ACKNOWLEDGMENTS

The authors wish to express their appreciation to the members of the Physics Department of Iowa State College for their advice and encouragement in carrying out this investigation. The authors also wish to thank Richard Barton for his work in preparing the borides and for the x-ray analysis of each sample.

Influence of Space Charge on the Potential Distribution in Mass Spectrometer Ion Sources

W. M. BRUBAKER
Consolidated Engineering Corporation, Pasadena, California
(Received October 11, 1954; revised manuscript received April 18, 1955)

The role of space charge as a factor influencing the potentials and potential gradients in a mass spectrometer ion source of the electron bombardment type is calculated. Planar equipotential surfaces are assumed, and the analysis then becomes that of a plane parallel positive ion diode. The "cathode" of the ion diode may be either emission-limited or space-charged-limited. The analysis considers the charge of the electrons in the ionizing sheet and the charge of the positive ions in the diode.

At a critical gas pressure the influences of the positive and negative charges on the potential of the ionizing region are equal and opposite, for small ionizing current. This concept leads to a pressure normalization in terms of the critical pressure. For a given source geometry and electron bombarding energy one can construct universal curves which give the potentials and the potential gradients as a function of the normalized gas pressure and the ratio of the ionizing electron current density to the repeller voltage. Experimental data are in agreement with the predictions of the theory.

INTRODUCTION

An appreciation of the phenomena which occur in a mass spectrometer ion source of the electron bombardment type is obtained from an analysis of what happens to the potentials and the potential gradients in the source when space charge effects are present. In spite of the small dimensions of such ion sources and the minuteness of the electron and ion currents used, space charges grossly alter the potential (and gradient) distribution. Thus the early portions of the ion trajectories may be sensitively influenced by space charge.

In this work the potential and the potential gradient at the ionizing region are calculated, and the results are presented as a set of general curves using normalized variables. The analytical expressions indicate the influence of ionizing electron current, gas pressure and ionizing probability, and the ion repeller voltage for any combination of these independent variables. The explicit dependence of space charge phenomena on the design parameters of the source geometry is obtained.

For purposes of analysis the equipotential surfaces in the source are assumed to be parallel planes, and any influence of the ion accelerating gradient which "leaks" through the first slit is ignored. Where possible the variables are normalized so that the analysis is carried out in dimensionless numbers.

The knowledge of the actual potential of the ionizing regions, as given in this paper, is of particular importance to one who does appearance potential studies.

GENERAL

The geometry of an idealized mass spectrometer ion source in which the equipotentials are assumed to be portions of parallel infinite planes is given in Fig. 1. The ionizing electron current is shown to flow in a sheet from the filament to the anode in the ionizing plane. Because the anode to which the electrons flow is of relatively small dimensions, it is shown as intercepting only a portion of the current in the large sheet.

The region between the ionizing plane and the first slit plane contains only positive ions, and constitutes a positive ion diode. The "cathode" of this diode is the ionizing plane, and the "anode" is the first slit plane. The gradient on the first slit side of the ionizing plane is frequently not zero, and in these instances the flow of current in the ion diode is emission-limited. The relation between the cathode gradient of an emission-limited plane-parallel diode and the current density is given in the literature.[1,2] The nomenclature and the results of reference 2 are used throughout this paper.

Typical potential distributions as functions of position between the repeller and the first slit plane are given in Fig. 2. In the absence of any space charge the voltage gradient in the source is uniform, and the potential distribution is given by curve 1. If an ionizing electron current I_1 is added at zero pressure (no positive ions) the electron space charge depresses the potentials as in curve 2. As we hold the ionizing current density constant and admit gas to the source in two consecutive steps the potentials rise as shown by curves 3 and 4. These potential distributions are based on the assumption that no negative ions are formed in the source.

In the course of this analysis it was discovered that it is extremely convenient to normalize the pressure in

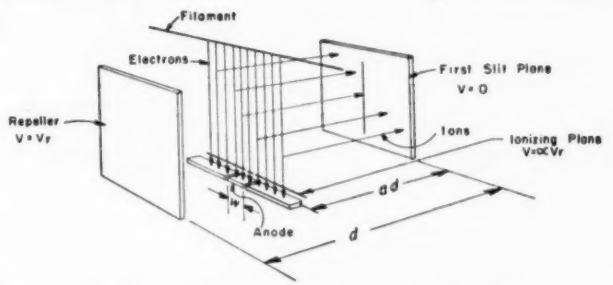


FIG. 1. Geometry of idealized ion source. The electron beam is constrained to the sheet by a magnetic field.

[1] H. F. Ivey, Phys. Rev. **76**, 554 (1949).
[2] W. M. Brubaker, Phys. Rev. **83**, 268 (1951).

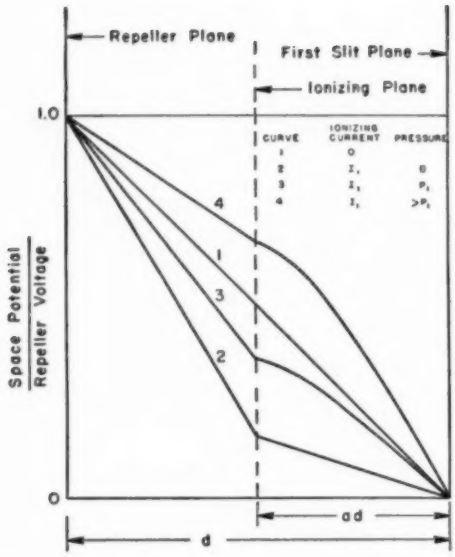


FIG. 2. Space potential *vs* position for typical conditions of operation. I_1 is an arbitrary value of ionizing electron current which is the same for curves 2, 3, and 4.

the source to that pressure at which, for the ionizing plane, the influence of the electron space charge in the ionizing sheet is exactly neutralized by the positive ion space charge at low ionizing currents. This pressure is noted as P_c, the critical pressure. A typical potential distribution at the critical pressure is shown in Fig. 3.

One of the main objectives is the calculation of the potential of the ionizing region as a function of the geometry, the ionizing current density, the pressure and the ionizing probability of the gas present in the source. This potential is normalized to the repeller potential by the dimensionless variable α. When α is known, several other parameters such as the gradient at the ionizing sheet, the ion current density relative to the space-charge-limited ion current density for the same diode potentials, etc., are also known. The potential of any plane in the source can be calculated with the equations given in reference 2, if desired.

The analysis is divided logically into three regions of increasing ion currents which are assumed to be proportional to the product of the anode current and the pressure. The regions are characterized as follows.

Region I

This region is bounded on the lower end by zero space charge, and on the upper end by the condition of space-charge-saturation (zero gradient) at the cathode of the positive ion diode. The single positive ion diode is emission-limited in this region.

Region II

Zero gradient on the first slit side of the cathode of the positive ion diode and a finite gradient on the repeller side define this region. Thus we have a single space-charge-limited diode.

Region III

At still higher ion currents the ions are formed at such a fast rate that they are not all drawn to the plane of the first slit. This happens when the potential of the ionizing sheet is equal to or greater than the repeller voltage. Thus the positive ion space charge causes the potential of the ionizing region to rise above that of the repeller, and we have zero gradient on both sides of the electron beam, or two space-charge-limited diodes in parallel.

The analysis of region I is made in two parts: one of a low current level, and one of a high current level. From the consideration of the low current level phenomena comes the concept of a "critical" pressure. This proves to be a most useful normalization which greatly aids the analysis at the high current levels.

The definitions of the symbols used are given in Table I. The variable i_V, the anode current in microamperes divided by the repeller voltage, is the only variable whose physical meaning is not obvious. It has the appearance of a transconductance, but the anode current and the repeller voltage are independent of each other. Their influence on the potentials in the source, however, is determined uniquely by their *ratio* independently of their individual magnitudes. This comes about because the influence of the electron beam is to depress the space potentials while the repeller voltage raises them.

In combination with the definitions of Table I, and the concept that the anode intercepts a portion of the electron current which flows in a sheet, i_V is expressed as:

$$i_V=\frac{wj_{el}\times 10^6}{V_r}, \tag{1}$$

$$i_V=\frac{w\sigma\times 10^6}{V_r}\left(\frac{2V_{el}e}{m_e}\right)^{\frac{1}{2}}. \tag{2}$$

Any dependence of V_{el} on the potential of the ionizing region is neglected. A form of Poisson's equation in rationalized mks units states

$$\Delta(dV/ds)=\sigma/\epsilon_0, \tag{3}$$

where in our case $\Delta(dV/ds)$ is the difference in the gradient on the two sides of the ionizing plane, and σ is the electron charge per square meter in the ionizing sheet.

The gradient on the first slit side of the ionizing plane is the cathode gradient of the positive ion diode. For emission-limited conditions the normalized density of ion current is related to the normalized cathode gradient by[2]

$$\rho=\tfrac{1}{2}+(\tfrac{1}{2}-\tfrac{3}{4}\gamma)(1+3\gamma)^{\frac{1}{2}}. \tag{4}$$

Combination of Eqs. (2) and (3) with the definition of

γ leads to

$$i_V=\frac{1}{A}\left[(1-\alpha)-\frac{1-a}{a}\alpha\gamma\right], \tag{5}$$

where

$$A=\frac{(1-a)d\times10^{-6}}{\epsilon_0 w}\left(\frac{m_e}{2V_{el}e}\right)^{\frac{1}{2}}$$

and is a constant for a given source operated at constant electron energy.

The flow of positive ions through the diode, according to the Child-Langmuir equation and the definition of ρ yields

$$j_{+ion}=\frac{4\rho}{9}\left(\frac{2e}{M}\right)^{\frac{1}{2}}\frac{(\alpha V_r)^{\frac{1}{2}}\epsilon_0}{(ad)^2}. \tag{6}$$

The physics of the ionization process gives us a relation between j_{+ion} and j_{el}

$$j_{+ion}=\frac{SP_{mm}V_r i_V}{w}\times10^{-6}. \tag{7}$$

The Eqs. (4) through (7) give us sufficient information to allow us to eliminate any three of the four variables α, γ, ρ, and j_{+ion}, and to express the remaining one as a function of i_V and the source parameters a, d, w, and V_{el} as given by A.

In region II the diode is space-charge-limited and so ρ and γ have the constant values of unity and zero, respectively. This greatly simplifies the equations.

Analysis, Region I

Low current solution.—The first instance in which approximation for ease of manipulation can be made without loss of accuracy at low levels is in Eq. (4).

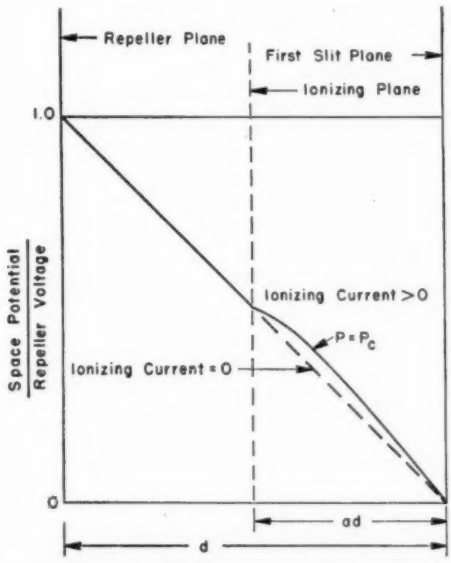


FIG. 3. Space potential *vs* position when P is the critical pressure, P_c. ($x=1$.)

TABLE I.

V_r = repeller voltage, volts, relative to the potential of the first slit plane, which is at $V=0$;
αV_r = potential of the ionizing plane;
V_{el} = energy of ionizing electrons, electron volts;
d = depth of ion source, repeller to first slit, meters;
ad = distance from ionizing region to the first slit, meters;
e = electronic charge, 1.6×10^{-19} coulombs;
j_{el} = electron current per meter in ionizing sheet;
M = mass of positive ion, kilograms;
m_e = mass of electron, 9×10^{-31} kilograms;
j_{+ion} = ion current density, amperes/sq meter;

$$i_V=\frac{\text{anode current in microamperes}}{\text{repeller voltage, volts}};$$

$$\gamma=\frac{\text{gradient at the cathode of the positive ion diode in the presence of positive ion space charge}}{\text{gradient at the cathode of the positive ion diode in the absence of positive ion space charge, at the same } \alpha};$$

$$\rho=\frac{\text{actual positive ion current density}}{\text{space-charge-limited positive ion current density, at the same } \alpha};$$

ϵ_0 = permittivity of free space, 8.85×10^{-12} farads/meter;
S = number of ion pairs formed per electron per meter of electron path at 1 mm Hg pressure;

$$x=\frac{\text{pressure, mm Hg}}{P_c\text{, the critical pressure, mm Hg}}=\frac{P_{mm}}{P_c};$$

w = width of ionizing electron sheet intercepted by the anode, meters.

For ρ small (4) becomes

$$\gamma\doteq1-\frac{16}{27}\rho. \tag{4a}$$

If we use Eqs. (4a), (5), (6), and (7) to eliminate ρ, j_{+ion} and γ from Eq. (5), we obtain

$$\alpha^{\frac{1}{2}}+a\alpha^{\frac{1}{2}}[Ai_V-1]=\tfrac{4}{3}a^2dSP_{mm}\left(\frac{MV_{el}}{m_eV_r}\right)^{\frac{1}{2}}Ai_V. \tag{8}$$

Equation (8) is readily solved for the conditions of small i_V by letting

$$\alpha=a(1+2\delta). \tag{9}$$

When Eq. (9) is set into Eq. (8).

$$\delta\doteq\frac{A}{2}\left[\tfrac{4}{3}dSP_{mm}\left(\frac{aV_{el}M}{V_rm_e}\right)^{\frac{1}{2}}-1\right]i_V. \tag{10}$$

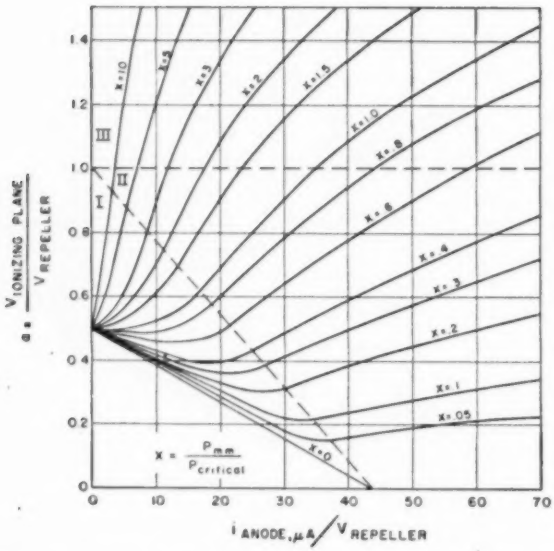


FIG. 4. Potential of the ionizing region as a function of i_V. regions I, II, and III are separated by the two dotted straight lines. The regions are characterized as: region I—one emission-limited diode; region II—one space-charge-limited diode; and region III—two space-charge-limited diodes in parallel.

Equation (10) is entirely satisfactory for the rate of change of δ (or α) with respect to i_V as $i_V \rightarrow 0$. It also shows the possibility of having the potential of the ionizing region independent of i_V. The pressure at which this occurs we shall call the critical pressure, P_c. It is given by

$$P_c = 3(V_r m_e / a V_{el} M)^{\frac{1}{2}} / 4dS, \tag{11}$$

and if we define x by the equation

$$P_{mm} = xP_c \tag{12}$$

Eq. (10) becomes

$$\delta \doteqdot (x-1)Ai_V/2 \tag{13}$$

and Eq. (9) becomes

$$\alpha \doteqdot a[1+A(x-1)i_V]. \tag{14}$$

For i_V small this is an excellent approximation.

The approximate solution as represented in Eqs. (13) and (14) is sufficiently accurate for many of the operating conditions found in actual practice. However, it is not applicable in all cases, and so it is expedient to consider the phenomenon without the approximations.

Large current solution.—If we combine Eq. (11) for the critical pressure with the Eqs. (4) through (7) we can calculate the potential variations and display the results for any given source in a set of curves in normalized variables. A combination of Eqs. (5), (6), (7), (11), and (12) yields

$$\frac{1+\dfrac{1-a}{a}\gamma}{\rho^{\frac{1}{2}}} = \frac{4}{9a}\left[\frac{2(1-a)}{A}\right]^{\frac{1}{2}}(i_V^{-\frac{1}{2}} - Ai_V^{\frac{1}{2}})\frac{1}{x^{\frac{1}{2}}} \tag{15}$$

or

$$F(\gamma,\rho) = x^{-\frac{1}{2}}G(i_V). \tag{16}$$

The explicit form of Eq. (16) is too complicated for direct computation. However, $F(\gamma,\rho)$ can be plotted as a function of either γ or ρ for a given source. Similarly, $G(i_V)$ can be plotted as a function of i_V. Then, using x as a parameter, families of curves can be obtained relating γ and ρ to i_V. Further, α as a function of i_V can be obtained from γ through the use of Eq. (5).

Analysis, Region II

In region II the positive ion current flows across the diode under space-charge-limited conditions at the ionizing sheet. The relation between the ion current density and the voltage across the diode is given by the Child-Langmuir equation

$$j_{+\text{ion}} = \frac{4}{9}\left(\frac{2e}{M}\right)^{\frac{1}{2}}\frac{(\alpha V_r)^{\frac{1}{2}}}{(ad)^2}\epsilon_0.$$

As the rate of ion production is uniquely determined by the anode current (of ionizing electrons), the ionizing probability, and the pressure of the gas in the source, it is the independent variable and the voltage the dependent variable. Thus, we can state

$$\alpha_{\text{region II}} = \alpha_{\text{boundary regions I and II}} \times \left(\frac{i_{V\text{region II}}}{i_{V\text{boundary regions I and II}}}\right)^{\frac{1}{2}}. \tag{17}$$

Analysis, Region III

In region III α is greater than unity. Thus, ions flow in both directions from the ionizing region as space-charge-limited current. Again, the rate of production of the ions is assumed to be entirely independent of the potential of the ionizing region. The potential of the region adjusts itself so that the sum of the two diode currents is equal to the rate of ion production.

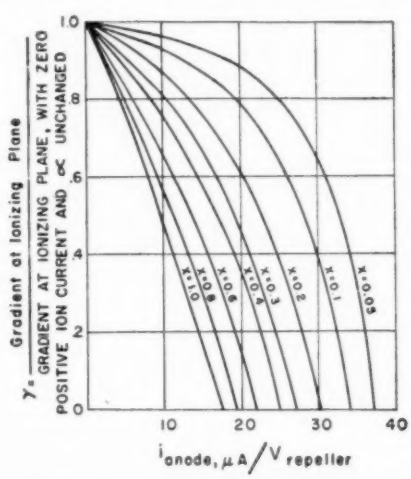


FIG. 5. Potential gradient at the ionizing plane as a function of i_V. The normalization is to the positive ion diode variables.

If we denote the ion current density to the first slit plane as j_1, and that to the repeller as j_2, we have

$$j_1=\frac{4}{9}\left(\frac{2e}{M}\right)^{\frac{1}{2}}(\alpha V_r)^{\frac{1}{2}}\epsilon_0/(ad)^2 \tag{18}$$

$$j_2=\frac{4}{9}\left(\frac{2e}{M}\right)^{\frac{1}{2}}[(\alpha-1)V_r]^{\frac{1}{2}}\epsilon_0/[(1-a)d]^2 \tag{19}$$

and

$$j_1+j_2=SP_{mm}j_{el}=\frac{4\left(\frac{2e}{M}\right)^{\frac{1}{2}}V_r^{\frac{1}{2}}\epsilon_0}{9d^2}\left[\frac{\alpha^{\frac{1}{2}}}{a^2}+\frac{(\alpha-1)^{\frac{1}{2}}}{(1-a)^2}\right]. \tag{20}$$

If we normalize the pressure as by Eqs. (11) and (12) this becomes

$$i_V=\frac{16\epsilon_0 w\times 10^6}{27xd}\left(\frac{2aeV_{el}}{m_e}\right)^{\frac{1}{2}}\left[\frac{\alpha^{\frac{1}{2}}}{a^2}+\frac{(\alpha-1)^{\frac{1}{2}}}{(1-a)^2}\right] \tag{21}$$

$$=B[\alpha^{\frac{1}{2}}+C(\alpha-1)^{\frac{1}{2}}]/x \tag{22}$$

where B and C are constants for a given source operated at constant electron ionizing energy.

APPLICATION OF THE THEORY

A set of universal curves has been made for a typical ion source of the electron bombardment type. The data for the source are given as follows: $a=0.5$, $d=0.1$ inch $=2.54\times10^{-3}$ meter, $w=0.05$ inch $=1.2\times10^{-3}$ meter, and $V_{el}=70$ volts.

The dependence of the potential of the ionizing plane upon i_V is shown in Fig. 4. The three regions are separated by the dotted straight lines. It is interesting to note that at low pressures the boundary between regions I and II is reached essentially as a result

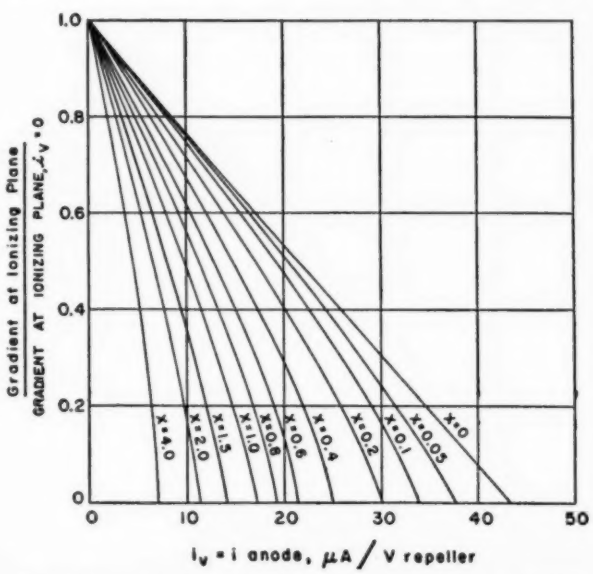


FIG. 6. Potential gradient at the ionizing plane as a function of i_V, normalized to the gradient when $i_V=0$.

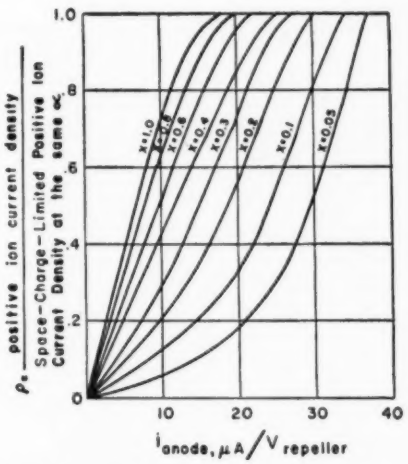


FIG. 7. Space-charge-saturation of positive ion current as a function of i_V, normalized to positive ion diode variables.

of the space charge of the electrons which lowers the potential across the positive ion diode to very low levels. At pressures above $x=1$ the potential of the ionizing region always increases with i_V, and at the higher pressures the potential of the ionizing region rises quite rapidly with i_V.

The normalized gradient on the first slit side of the ionizing region is shown in Fig. 5. γ is defined in terms of the positive ion diode variables. Thus the gradient to which the data of Fig. 5 are normalized is the gradient which would exist if the potential of the ionizing region remained constant while the pressure was reduced to zero.

Because the normalization of γ is to a gradient which is not experimentally realizable, it is interesting to plot the actual gradient normalized to its value at $i_V=0$. To do this we multiply by α/a. This is done in Fig. 6. Here we see the influence of space charge on the gradient at the ionizing plane. It is obvious that this gradient is decreased by both the electron and the ionic space charges.

The normalized positive ion current density, ρ, as a function of i_V, is given in Fig. 7. Here again we see that at low pressures space-charge-limited conditions are reached primarily because the electron space charge lowers the diode voltage. At the higher pressures the rate of approach to space-charge-limited conditions is slowed by an increase of the diode voltage caused by the positive ion space charge. The normalization is in terms of the diode variables.

EVALUATION OF THE APPROXIMATIONS

In order to facilitate a mathematical statement of the problem, two simplifying assumptions were made: (A) All equipotentials are parallel planes; and (B) the ionizing sheet of electrons is of infinitesimal thickness.

Neither of these is realized in practice. Assumption (A) makes the greater departure from reality, as the dimension of the electron beam in the direction parallel

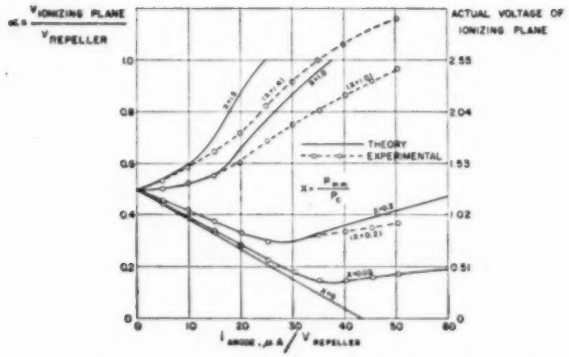


FIG. 8. Comparison of theory and experiment. The anode current was empirically adjusted to fit the low pressure curve, $x=0.05$. Remaining data were then plotted without further adjustment.

to the repeller plane is quite small. This causes the capacity of the ionizing region to be much greater than it otherwise would be. In addition the warping of the equipotentials in the vicinity of the space charge may be expected to have an important influence on the optical properties of the source. This may well be the mechanism of the influence of the presence of a second gas on the sensitivity of the instrument to a first gas.

Assumption (B) is of lesser importance. The influence of space charge on potentials outside the volume in which the charges reside is essentially a function of the total charge involved and independent of the distribution of the charges. Thus outside the ionizing region itself the potentials are relatively unaffected by the thickness of the electron beam. This paper ignores the distribution of potentials and gradients within the ionizing sheet, but it is felt that the conclusions drawn are little influenced by the distribution of electron current in the direction perpendicular to the plane of ionization.

EXPERIMENTAL VERIFICATION OF THE THEORY

The soundness of the analysis is supported by experimental data which relates the potential of the ionizing region with anode current, repeller voltage, and the pressure. The data shown in Fig. 8 were taken by observing the change in ion accelerating voltage required to keep the ion beam on the collector of the mass spectrometer as the anode current and the pressure were varied. The actual potentials of the ionizing region are given as well as the relative values in terms of α.

The anode current as plotted was normalized to fit the theoretical curve at the lowest pressure, and then the rest of the data were taken without further normalization. This empirical normalization is required because the electron beam is not in the form of a very broad (infinite) sheet, as was assumed for purposes of calculation. The capacity per unit area of the ionizing sheet is greater because of its finite size.

Since the influence of the electron space charge on the potential of the ionizing region varies inversely as the electrostatic capacity of the beam to its environs, we would expect the dependence of the space potential on the electron beam to be less than that calculated for infinite sheets. The same argument applies to the ions. In Fig. 8 the actual anode currents have been reduced by an arbitrary factor of 1.65 to obtain the degree of agreement shown between experiment and theory.

CONCLUSIONS

The essential characteristics of the potentials and gradients in a mass spectrometer ion source are derived by treating the ion accelerating region as an emission-limited or space-charge-limited diode as the conditions require. A comparison of the theory with experimental data shows that the theory is correct in its essentials. When allowance is made for the increased capacity of the ionizing region owing to its finite size, the agreement between experiment and theory is satisfactory.

The knowledge of the dependence of the source potentials on the source variables (ionizing current, repeller voltage, pressure, etc.) is vital for appearance potential investigations. It is also necessary for an understanding of the influence of space charge on the optical properties of the source.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the benefit of discussing this problem with Dr. Clifford Berry, Dr. Harold W. Washburn, and Mr. Lawrence Hall. Mr. Hall also obtained the experimental data.

Slowing Down Distribution of U^{235} Fission Neutrons from a Point Source in Light Water*

J. E. Hill,† L. D. Roberts, and T. E. Fitch
Oak Ridge National Laboratory, Oak Ridge, Tennessee
(Received February 8, 1955)

The results of a measurement of the slowing down distribution in water of neutrons from a point fission source to indium resonance energy, 1.458 ev, are given. The second, fourth, sixth, and eighth moments of the measured distribution are calculated and have the values, $\bar{r}^2=184.7$ cm^2, $\bar{r}^4=1.222\times10^5$ cm^4, $\bar{r}^6=2.27\times10^8$ cm^6, and $\bar{r}^8=8.6\times10^{11}$ cm^8.

INTRODUCTION

PREVIOUS measurements of the age, i.e., the second moment of the slowing down distribution,[1] of fission neutrons from a point source in water, to indium resonance energy,[2] 1.458 ev, have been made by Fermi, Anderson, and Nagle.[3] In this earlier work, the neutron distribution was measured only to a distance of about 30 cm from a "point" source; the neutron sources were rather large, making geometry corrections somewhat uncertain; and the water tank used was perhaps smaller than the optimum size. In the measurements reported here, the indium resonance neutron distribution has been measured out to 92.20 cm. At the expense of using a rather large number of neutron source sizes, the geometry corrections were kept small. The water tank which we used held a five-foot cube of water, which is equivalent to an infinite body of water in this experiment.

In some measurements[4] of the neutron slowing down distribution in graphite, we found that indium foils were appreciably activated at neutron energies above 1.458 ev. As will be discussed later, it was found through the use of B^{10} covered indium foils that this higher energy activation does not appreciably affect the approximately[5] 1.458 ev neutron distribution, and that the moments calculated from the neutron distribution measured with cadmium covered indium foils are probably very close to the correct ones for neutrons of energy near 1.458 ev. Since the neutron distribution was measured to 92.20 cm from the source, it was possible to calculate the moments of the distribution $\bar{r}^2$, $\bar{r}^4$, $\bar{r}^6$, and $\bar{r}^8$. (See Table I.) Between 0 cm and 92.20 cm from the point source the indium resonance neutron flux was found to change by a factor of 4.27×10^{-8}.

PROCEDURE

Briefly, the technique of the experiment was to place a large tank of water on the thermal column of the Oak Ridge National Laboratory graphite reactor. Thermal neutrons coming up from the pile entered the tank through the bottom. A fast neutron source plate of U^{235}–Al alloy was suitably located relative to the bottom of the tank. Some of the entering thermals were captured by this source plate thus generating fission neutrons. The slowing down distribution of these neutrons was measured with indium foils covered by aluminum, cadmium, or B^{10}.

DETAILS OF THE APPARATUS

The water tank mentioned above was 5 ft by 5 ft by 6 ft high and was fabricated from $\frac{1}{2}$ in.-thick aluminum plate. An "I" beam with a 3 in. web was mounted across the top and center of the tank and on this was mounted a set of guides. The neutron source plate and the indium foils in suitable boxes were supported in the water in either of two 2-S aluminum frames, Figs. 1 and 2, which were in turn supported at the end of a 1 in. o.d. aluminum tube.

When this tube was inserted in the previously mentioned guides with neutron source and foils immersed in the water the neutron source plate and foils lay in a plane parallel to the bottom of the tank and on the central vertical axis of the tank. Due to the precision construction of this guide and of the framework holding the source plate and the indium foils, all spacings of the source and foils from each other or from the bottom of the tank were accurate to better than 0.03 cm.

Figure 1 shows the lower end of a tube 8 ft long. The upper end was inserted in the guide already mentioned which was used to locate the mechanism in

TABLE I. The moments $\bar{r}^n$ of the distribution of about 1.458 ev neutron from a point fission source in water.

n	$I_n=\int_0^\infty A_s r^n dr$	Fraction of I_n extrapolated	$\bar{r}^n=\dfrac{\int_0^\infty A_s r^{n+2} dr}{\int_0^\infty A_s r^2 dr}$
2	2.092×10^6	1.40×10^{-5}	$\bar{r}^2=184.7$ cm^2
4	3.863×10^8	7.95×10^{-4}	$\bar{r}^4=1.222\times10^5$ cm^4
6	2.556×10^{11}	0.0130	$\bar{r}^6=2.27\times10^8$ cm^6
8	4.750×10^{14}	0.0857	$\bar{r}^8=8.6\times10^{11}$ cm^8
10	1.809×10^{18}	0.333	

* Previously reported in the U. S. Atomic Energy Commission report ORNL 181.

† Now with The Rand Corporation, Santa Monica, California.

[1] See, for example, Glasstone and Edlund, *The Elements of Nuclear Reactor Theory* (D. Van Nostrand Company, Inc., New York, 1952).

[2] V. L. Sailor, Phys. Rev. **87**, 222 (1952).

[3] Fermi, Anderson, and Nagle, Argonne National Laboratory report CP-1531(A2209), (unpublished).

[4] Roberts, Hill, and McCammon, Phys. Rev. **80**, 6 (1950).

[5] We refer to *approximately* 1.458 ev neutron energies because of self-protection effects in the indium foil activation, and because the activation in the wings of the line is not symmetrical. For intensity reasons, it is impractical to use an indium foil thin in the sense of neutron absorption at the resonance energy.

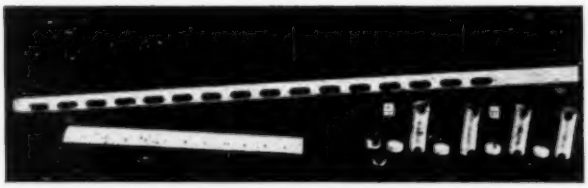


FIG. 1. Indium foil covers and the supporting mechanism for measurements close to the source.

the water tank. The lower end held the foils in cadmium boxes as follows. The foils were placed in boxes *B* which closely fit in the tube. Then the spacers *S* and the boxes *B* were inserted alternately into the tube. Eight boxes were put in for each run, the foils being 3.00 in. apart, and the first foil being either very nearly 3 in. or approximately 1.5 in. from the neutron source. The oblong holes shown in the tube, and the notches at each end of the spacers, were cut to permit water to freely fill the tube. The assembly of boxes and spacers was held in place in the tube by suitable stops. The fission neutron source plate was attached by screws to the lower end of the tube in an accurately reproducible fashion.

Figure 2 shows the aluminum holder used for the large size cadmium box *A* and B^{10} filled box *C*. (Also, see Table II.) One of the three long slotted aluminum tubes was removable to facilitate the loading and unloading of the cadmium (or boron) boxes in the aluminum frame. The measurements were made using sets of four foils per run. In a given run the foils were always 3.00 in. apart, as shown in the photograph. When either the 2 cm or 5 cm diameter disk source was used with this frame, it was attached to the bottom of plate *E* by a screw. When the 10 cm square source plate was used, it was inserted into the slot at the left side of holder *D*, and the latter was snapped onto plate *E* using a spring latch mechanism.

Two sizes of cadmium boxes, *A* and *B* were used. Box *A*, Fig. 2 was rectangular in shape and sufficiently large to hold a foil 4 cm×6.35 cm. The wall thickness of this box was 2.74 g/cm² of cadmium. The other cadmium box *B* Fig. 1 was of pill box shape with an internal diameter of 1.52 cm, sufficient to accept foils 1-cm square. The wall thickness was 3.07 g/cm². The boxes were clamped shut with a small amount of grease around the edges to keep water out during a foil activation. To make measurements at large distances from the neutron source, the indium foils were covered with 0.020 in. thick aluminum sheet rather than with the cadmium boxes.

In addition to the cadmium and aluminum foil covers discussed in the foregoing, two types of boxes were made to cover foils with B^{10} powder. The B^{10} powder was used without any binding agent to avoid neutron slowing down effects which a binding material might introduce. In one set up an aluminum box, *C*, with a steel cover was made. This box was large enough to accept a 4 cm×6.35 cm indium foil plus suitable layers of B^{10}. The other B^{10} box consisted of a set of brass shells of pill box proportions. In appearance they were very similar to cadmium box *B*, Fig. 1. However, both the top and bottom were hollowed out and filled with the B^{10} powder. In box *C*, there was 0.245 g/cm² of boron and in the pill box, there was 0.337 g/cm². This boron was 88% boron ten by weight. Both assemblies were sealed against water.

The indium foils used were approximately 0.10 g/cm² thick and within the measurement of a given section of the distribution, the foil weights were within 0.5% of the average weight. For different sections of the curve where the average weight of the foils differed slightly, the sections were normalized together, usually by one or more common points. The area of the foils used is given in Table II.

The different size sources of the fission neutrons were plates of aluminum-uranium 235 alloy. The uranium was 96% U^{235} and the alloy 18% uranium by weight (eutectic composition). The alloy plate neutron sources were 0.080 in.±0.004 in. thick and of various areas as specified in Table II and as described in the following. This thickness of uranium, with the surrounding water reflector, was sufficient to absorb about 95% of the incident thermal neutrons from the thermal column.

MEASUREMENT PROCEDURE

In general, the indium foils were activated for an hour and counted for an hour. Two counters were used, and two foils were counted simultaneously. The ratio between the two counters was obtained using a single strongly activated indium foil and this ratio was used in all of the calculations. In addition, the ratio

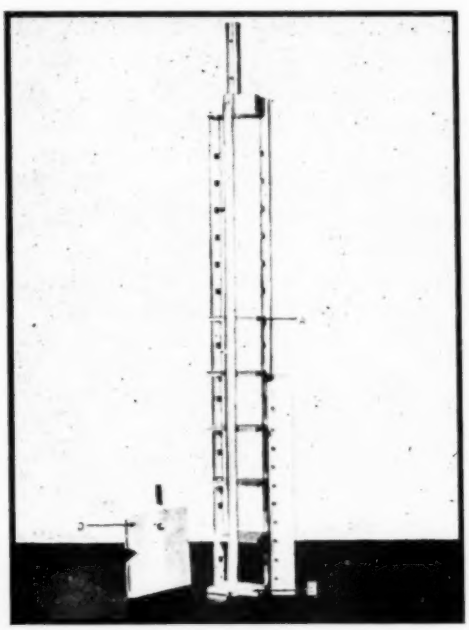


FIG. 2. Indium foil covers and the supporting framework for measurements far from the source.

between the two counters was checked daily for constancy using a uranium oxide standard.

The side of the foil nearest to the neutron source was counted in all cases. In addition, the ratio of the activities of the two sides of the foil was measured at a number of selected positions and this ratio was plotted as a function of the distance from the neutron source. Then, using this graph, all of the measured points were corrected to give the average activity of the two sides. In all of the activities reported here, it is this average which is given.

The activity which one would like to measure would be that for neutrons coming from an infinitesimal source and being detected by an infinitesimal detector. Actually, the source and detector sizes must be fairly large to have a measurable foil activity, and as a result, one must correct this measured activity to the desired activity density at the source-foil axis. In practice this correction was rarely as much as ten percent and was usually much less. The formulas used are straightforward, and only the general method of derivation is outlined in the Appendix. All activities A_s quoted are these activity densities.

Even with the largest neutron source, 30.5 cm×30.5 cm square and with the larger size foils, the foil counting rate became unacceptably small at about 70 cm for cadmium covered foils. To obtain the distribution between about 70 cm and 92.20 cm, measurements were made using indium foils in aluminum covers, Table II. The aluminum covers were used to provide mechanical support for the foils and to prevent contamination.

TABLE II. Summary of source size and location, foil size and foil box material, aluminum supporting framework used, and range of r over which the given experimental condition applied.

Range of distances from source r	Neutron source geometry	Neutron foil size and foil box	Aluminum frame used
2.35 cm to 15.02 cm	Disk 5.08 cm diam 10 cm from bottom of water tank	Disk foil 1.43 cm diameter in Cd box *B*, Fig. 1	As shown in Fig. 1
12.06 cm to 22.00 cm	Disk 1 cm diam in contact with bottom of water tank	1 cm square foil in Cd box *B*, Fig. 1	As shown in Fig. 1
15.32 cm to 41.99 cm	Disk 2 cm diam in contact with bottom of water tank	Rectangular foil 4 cm×6.35 cm in Cd box *A*, Fig. 2	As shown in Fig. 2
34.54 cm to 61.21 cm	Square 10 cm×10 cm in contact with bottom of water tank	Rectangular foil 4 cm×6.35 cm in Cd box *A*, Fig. 2	As shown in Fig. 2
54.10 cm to 79.96 cm	Square 30.5 cm ×30.5 cm in contact with bottom of tank	Rectangular foil 4 cm×6.35 cm in Cd box *A*, Fig. 2	As shown in Fig. 2
69.34 cm to 92.20 cm	Square 30.5 cm ×30.5 cm in contact with the bottom of water tank	Rectangular foil 4 cm×6.35 cm in aluminum cover (not shown)	As shown in Fig. 2
2.28 cm to 7.54 cm	Disk 5.08 cm diam 10 cm from bottom of tank	1 cm square foil covered with 0.337 g/cm² boron	As shown in Fig. 1
7.54 cm to 17.60 cm	Disk 5.08 cm diam on bottom of tank	1 cm square foil covered with 0.337 g/cm² boron	As shown in Fig. 1
17.60 cm to 35.59 cm	Square 10 cm×10 cm on bottom of tank	Rectangular foil 4 cm×6.35 cm in boron cover 0.245 g/cm²	As shown in Fig. 2

TABLE III. Distribution of indium resonance activity from a fission source in light water (with transport and geometry corrections).

r(cm)	A_s	r(cm)	A_s
2.35	9167.	26.75	16.64
3.62	7274.	30.56	7.126
6.15	4281.	30.73	7.036
7.42	2994.	34.54	3.616
9.95	1445.	38.38	1.727
11.22	1008.	42.16	872.3×10^{-3}
12.07	775.8	45.72	445.4×10^{-3}
12.70	614.5	45.97	464.4×10^{-3}
13.33	521.5	49.78	227.8×10^{-3}
13.75	460.4	53.59	116.6×10^{-3}
14.61	354.3	54.10	108.4×10^{-3}
15.02	318.6	57.40	661.2×10^{-4}
15.24	292.8	57.91	610.4×10^{-4}
15.49	273.7	61.21	381.6×10^{-4}
15.88	255.9	61.72	354.3×10^{-4}
17.15	176.9	65.55	184.3×10^{-4}
18.42	126.2	69.34	100.4×10^{-4}
19.05	107.0	73.15	59.19×10^{-4}
19.69	88.17	76.96	35.91×10^{-4}
20.32	77.10	80.77	20.28×10^{-4}
20.96	65.90	84.58	11.27×10^{-4}
22.23	46.83	88.39	7.529×10^{-4}
22.94	40.32	92.20	4.439×10^{-4}
23.11	37.77		

The justification for this procedure rests on the work of Rush,[6] and on some of our measurements in which it was found that, at large distances from the neutron source, a distribution measured with cadmium covered indium foils and a second one measured with aluminum covered foils were proportional within experimental error. The range of r over which we checked this proportionality was about 60 cm to about 76 cm and in the work of Rush, 10 cm to 40 cm. This proportionality over a moderate range of r is not so surprising when one remembers that, due to the very large scattering cross section of hydrogen, the age from 1.458 ev to thermal is only of the order of one cm² while the age from fission energies to 1.458 ev is 30.8 cm². Thus the first moment of the neutron slowing down distribution from fission energies to thermal energy is only slightly greater than that to 1.458 ev, leading one to anticipate that the distributions will be very similar. This may be visualized physically in the following way: At the high initial source energies of the neutrons, the hydrogen cross section is small, and the neutrons may move out far from the source before they make a collision in which a great deal of energy is lost. After this collision, if the neutron energy is below say a few hundred kilovolts, the corresponding hydrogen cross section will be relatively quite large and the neutron will not be able to diffuse far from this point of first collision. Thus, to a first approximation, the slowing down distribution in water far from the source will be given by a distribution of first collisions. To our knowledge, this treatment was first used by Fermi.

[6] J. H. Rush, Phys. Rev. **73**, 271 (1948).

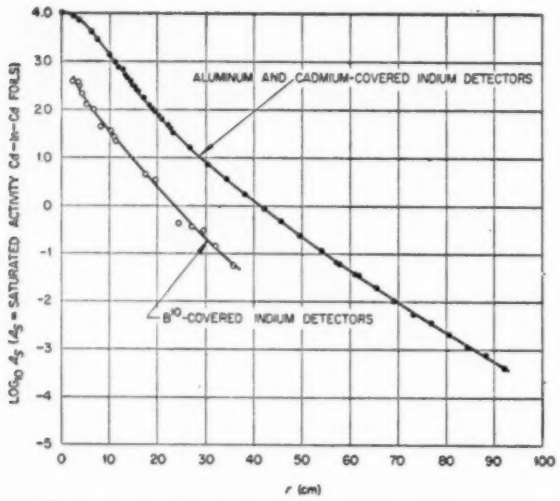


FIG. 3. Graph of the measured indium resonance energy neutron slowing down density.

BOUNDARY CONDITIONS

The best boundary conditions for a neutron slowing-down measurement would be to have the neutron source located at the center of a very large volume of the slowing-down medium. When a fission neutron source is used, however, the thermal neutrons from the thermal column which are used to produce the fissions have to diffuse through this slowing-down medium. If the neutrons have to diffuse very far, their number will be greatly attenuated, thus weakening the fission source. In water, the thermal neutron relaxation length is roughly one inch and a fission source placed four inches in from the water boundary will give less than 2% of the number of fission neutrons that the same source would give at the water surface. For measurements at large distances from the source, this decrease of source strength would be a serious handicap. A scheme to overcome this difficulty which seemed reasonable was to make measurements close to the neutron source with the source 10 cm in from the water boundary, and to make the measurements at large distances from the source with the source at the water boundary, i.e., at the bottom of the tank, Table II. To check up on this procedure, we measured two distributions in the range zero to twenty centimeters from the source: (1) with the source at the water boundary and (2) with the source 10 cm into the water. Close to the source the curves differed in shape, as one would expect, but beyond ten cm the two curves were proportional. This indicated that it would be satisfactory to make measurements at distances greater than ten cm from the source with the source at the water boundary. Since ten cm is many transport mean free paths for indium resonance neutrons, and on the basis of the above experiment, we felt that correct results would be obtained in the region from zero to ten cm with the neutron source ten cm from the water boundary.

DATA, CALCULATIONS, AND DISCUSSION

Using the equipment and methods described above, the distribution in water of indium resonance neutrons originating at a point fission source was measured in the range of distances from the source, $0 \leq r \leq 92.20$ cm. This distribution is given in Table III and Fig. 3. When this work was completed, we were not sure to what neutron energy the curve corresponded because we knew from other experiments[4] that the indium foils were activated to some extent by neutrons in the kilovolt region.

To determine the extent of this higher energy neutron activation and to correct for it, measurements were made using B^{10} covered foils, Table II. The amount of boron used was sufficient to remove neutrons in the energy range below 10 ev quite effectively. The results of these measurements are given in Table IV and in the graph, Fig. 3. From Fig. 3, it is seen that the boron covered curve is about 3% of the cadmium covered curve, and also that the two curves are almost proportional over a considerable region, though the statistics of the "boron covered" curve are rather poor. This proportionality of the distribution measured under boron to that measured under cadmium may be understood in terms of a distribution of first collisions, as discussed in the foregoing. The second moment $\bar{r}^2$ of the two curves was found to be almost the same, i.e., within about 10%. Because of the small magnitude of the boron covered indium foil activity, and because of the near proportionality of the two curves, the correct moments for an approximately 1.458 ev detection energy will be obtained within statistical error directly from the cadmium covered indium distribution without a correction for the contribution by higher neutron energies.

To calculate the moment of the measured distribution, $\bar{r}^n$, the experimental data were plotted on a $\log_{10} A_s \cdot r^2$ *vs* r plot and the best smooth curve was drawn through the data. Then points were read from this curve every cm and these smoothed points were used in all of the numerical integrations. The smoothed points are given in Table V. We define the integrals,

$$I_n = \int_0^\infty A_s r^n dr. \tag{1}$$

TABLE IV. A_s *vs* r for indium covered with boron.

r(cm)	A_s(counts/min)	r(cm)	A_s(counts/min)
2.28	399.7	10.93	27.1
2.37	397.1	11.35	22.4
3.65	377.3	17.60	4.56
3.78	303.6	19.56	3.42
4.43	210.2	24.64	0.412
5.29	129.4	27.17	0.357
6.73	98.7	29.71	0.292
8.24	44.9	32.25	0.141
10.23	36.3	35.59	0.0592

Then in a spherical geometry the nth moment $\bar{r}^n$ is given by the ratio

$$\bar{r}^n = I_{n+2}/I_2. \tag{2}$$

The portions of the I_n integrals from $r=92$ cm to $r=\infty$ were obtained analytically by using an extrapolating function based on the "distribution of first collisions" concept. In this region, we assumed

$$A_s = ke^{-r/\lambda}/r^2 \tag{3}$$

with $\lambda=8.786$ cm and $k=1.324\times10^5$ which are the values obtained from the experimental curve near 92 cm. The general form of these integrals is

$$k\int_{r_0}^{\infty} e^{-r/\lambda} r^n dr = k\lambda e^{-r_0/\lambda}\sum_{i=0}^{n}[i!\ {}_nC_i r_0^{(n-i)}\cdot\lambda^i], \tag{4}$$

where

$$n=0,2,4\cdots \text{ and } {}_nC_i=\frac{n!}{i!(n-i)!}.$$

The values of I_n and the moments, $\bar{r}^n$, calculated from the distribution are given in Table I. Table I also gives the fraction of each I_n integral contributed by the extrapolated part from $r=92$ cm to $r=\infty$.

In comparing our results with the earlier work of Fermi *et al.*, it is seen that our value of $\bar{r}^2=184.7$ cm^2 is rather less than the earlier value of 194 cm^2. The source of this difference may be seen in a comparison of the earlier distribution with ours. It is found that in the region from 10 cm to 24 cm, the earlier work and our measurements are in excellent agreement. At smaller and at larger r, however, the earlier measurements fall higher than ours, probably due to geometry close to the source and to statistics far out. The effect of this was that much too large a relaxation length λ was used in extrapolating the earlier distribution to infinity. Since the 194 cm^2 was about 30% extrapolated, the difference between the earlier work and our results can readily be explained in this way.

Reference to Table I shows that an appreciable part of $\bar{r}^6$ and $\bar{r}^8$ is extrapolated. The values given are probably underestimates. The reason for this lies in the fact that the extrapolating function used is that of a distribution of first collisions with constant relaxation length λ, whereas it is clear that λ is not a constant but gradually increases with r, due to the hardening of the neutrons.

Statistically, these data are of high accuracy. The majority of the measured activities, Table III, represent over 50 000 counts and no point represents fewer than 10 000 counts. Thus, statistically, the measured portion of the moments, Table I, should be good to better than one percent. Errors from other sources are estimated to be of similar magnitude.

TABLE V. Slowing down distribution of fission neutrons to indium resonance. A_s is saturated activity in indium under cadmium. Values in ( ) were extrapolated. r is in (cm).

r	A_s	r	A_s	r	A_s	r	A_s
(0)	(10 400)						
(1)	(10 190)	25	24.27	49	0.2661	73	0.6026×10^{-3}
(2)	(9 441)	26	19.28	50	0.2265	74	0.5248×10^{-3}
3	8 497	27	15.49	51	0.1905	75	0.4571×10^{-3}
4	7 162	28	12.59	52	0.1622	76	0.3981×10^{-3}
5	5 755	29	10.23	53	0.1380	77	0.3507×10^{-3}
6	4 519	30	8.318	54	0.1161	78	0.3020×10^{-3}
7	3 350	31	6.839	55	9.886×10^{-2}	79	0.2630×10^{-3}
8	2 512	32	5.625	56	8.414×10^{-2}	80	0.2291×10^{-3}
9	1 862	33	4.624	57	7.161×10^{-2}	81	0.2022×10^{-3}
10	1 396	34	3.856	58	6.095×10^{-2}	82	0.1748×10^{-3}
11	1 047	35	3.199	59	5.188×10^{-2}	83	0.1514×10^{-3}
12	776.2	36	2.661	60	4.416×10^{-2}	84	0.1318×10^{-3}
13	582.1	37	2.213	61	3.758×10^{-2}	85	0.1150×10^{-3}
14	431.5	38	1.841	62	3.199×10^{-2}	86	0.1000×10^{-3}
15	323.6	39	1.531	63	2.754×10^{-2}	87	0.8710×10^{-4}
16	242.7	40	1.303	64	2.344×10^{-2}	88	0.7586×10^{-4}
17	182.0	41	1.048	65	1.995×10^{-2}	89	0.6683×10^{-4}
18	138.0	42	0.9120	66	1.718×10^{-2}	90	0.5821×10^{-4}
19	105.9	43	0.7586	67	1.479×10^{-2}	91	0.5070×10^{-4}
20	81.28	44	0.6383	68	1.274×10^{-2}	92	0.4416×10^{-4}
21	63.10	45	0.5270	69	1.096×10^{-2}		
22	49.55	46	0.4519	70	0.9441×10^{-3}		
23	38.90	47	0.3758	71	0.8128×10^{-3}		
24	30.55	48	0.3162	72	0.6998×10^{-3}		

ACKNOWLEDGMENTS

The authors wish to express their appreciation to Dr. A. M. Weinberg for many interesting and helpful discussions of this work and to Mrs. Grace McCammon Estabrook for help in the measurements.

APPENDIX

Geometry Corrections

In this experiment all of the sources and detecting foils were thin plane configurations. They were arranged with their planes accurately perpendicular to the axis of the distribution, and with their geometrical centers on this axis. Let r_0 be the distance between the center of the source and that of a given foil, and let r be the variable distance between arbitrarily placed elements of area dS and dS' on the source and detecting foil whose total areas were S and S', respectively. Let $A_s(r)/SS'$ be the activity induced in a unit area of the detecting foil by a unit area of source when the separation of these areas is r. If $A_m(r_0)$ is the measured activity, then

$$A_m(r_0)=\int_S\int_{S'}\frac{A_s(r)}{SS'}dS'dS. \tag{5}$$

If the dimensions of the foil and source are small compared to r_0 then one may approximate the expression for r in terms of r_0 and the coordinates of the area elements dS and dS' by the first two terms of a binomial series expansion. Then, using Eq. (5) and the measured distribution $A_m(r_0)$ we calculated the slowing down distribution density, $A_s(r_0)$, Table III.

Slowing Down Distribution to Indium Resonance of U^{235} Fission Neutrons from a Point Fission Source in Two Aluminum Light Water Mixtures

L. D. Roberts, J. E. Hill,* and T. E. Fitch
Oak Ridge National Laboratory, Oak Ridge, Tennessee
(Received February 8, 1955)

The mean square slowing down length, $\bar{r}^2$, has been measured for two aluminum light water mixtures, the aluminum-to-water volume ratios being 1:1 and 1:2 by volume. The values of $\bar{r}^2$ obtained are 460.7 cm^2 and 297.4 cm^2, respectively.

INTRODUCTION

SINCE aluminum is an important structural material in water-moderated reactors, it was considered important to determine the mean square slowing down length, $\bar{r}^2$, in light water aluminum mixtures. We have measured this quantity for two aluminum-water mixtures, the volume ratios being 1:1 and 1:2 for the two samples. The values of $\bar{r}^2$ obtained for the 1:1 mixture was 460.7 cm^2 and for two volumes of water to one volume of aluminum, $\bar{r}^2 = 297.4$ cm^2.

APPARATUS AND MEASUREMENT PROCEDURE

A 2-S aluminum tank 5 ft by 5 ft by 6 ft high, resting on the thermal column at the top of the Oak Ridge National Laboratory graphite reactor was filled with a stacked grid structure of 2-S aluminum plates of thickness 0.250 in.$\pm$0.003 in. To obtain the two mixtures studied, these plates were separated by one-inch diameter aluminum spacers of thickness 0.250 in. $\pm$0.003 in. for the 1:1 mixture and of 0.500 in.$\pm$0.006 in. thickness for the 1:2 mixture. The plates lay in horizontal planes parallel to the top of the pile and perpendicular to the central vertical axis of the tank along which the measuring foils were located. After the grid structure had been assembled in the tank, the latter was filled with water, particular care being taken to insure the removal of air bubbles.

In all of the measurements reported here, indium foils 4 cm by 6.35 cm by about 0.10 g/cm^2 thick were used. They were of the same thickness to within 0.5 percent. The foils were enclosed in cadmium boxes with a wall thickness of 0.320 cm or in thin aluminum covers, and they lay in a plane parallel to the plates of the aforementioned aluminum grid structure. The aluminum covers were used for $r > 52.7$ cm for the 1:1 mixture and $r > 46.8$ cm for the 1:2 mixture. See the previous paper (referred to as paper I) for a discussion of this point. A cylindrical structure of aluminum disks 0.250 in. thick and 4 in. in diameter, spaced the same as the aluminum grid for a given mixture, was used to support the cadmium and aluminum covered foils, and this structure was so arranged that it could be easily withdrawn from the tank to insert and remove the foils. Figure 1 shows the top of the water tank and the upper surface of the aluminum grid with the above cylindrical structure partially withdrawn. This structure was quite precisely made so that all of the values of r, the distance from the foil to the neutron source, were good to at least 0.03 cm. The source of the fission neutrons which was one of those used for the "age in water" measurement, paper I, was a disk of U^{235}-Al alloy 5.08 cm in diameter and 0.2 cm thick. The alloy was of eutectic composition, or 18% uranium which was 96% U^{235}. The source was mounted on the aforementioned cylindrical structure on the lower end and adjacent to the bottom of the tank for values of r greater than 22.86 cm for the 1:1 mixture and 16.31 cm for the 1:2 mixture. For measurements in the region $0 < r \leqslant 22.86$ cm for the 1:1 mixture, the source was located at 20 cm from the boundary of the Al-H_2O cube; and for the 1:2 mixture, with values $0 < r \leqslant 16.31$ cm, the source

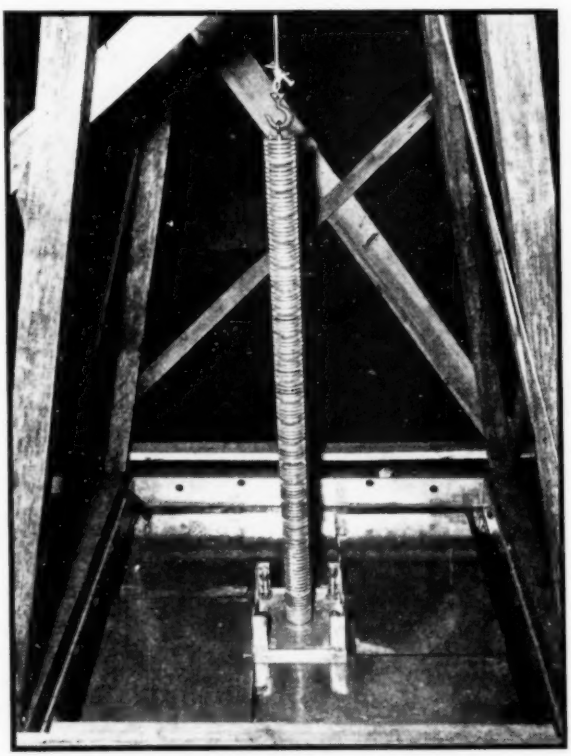


FIG. 1. A view of the top of the water tank showing the upper surface of the aluminum grid with the foil supporting structure partially withdrawn.

* Now with The Rand Corporation, Santa Monica, California.

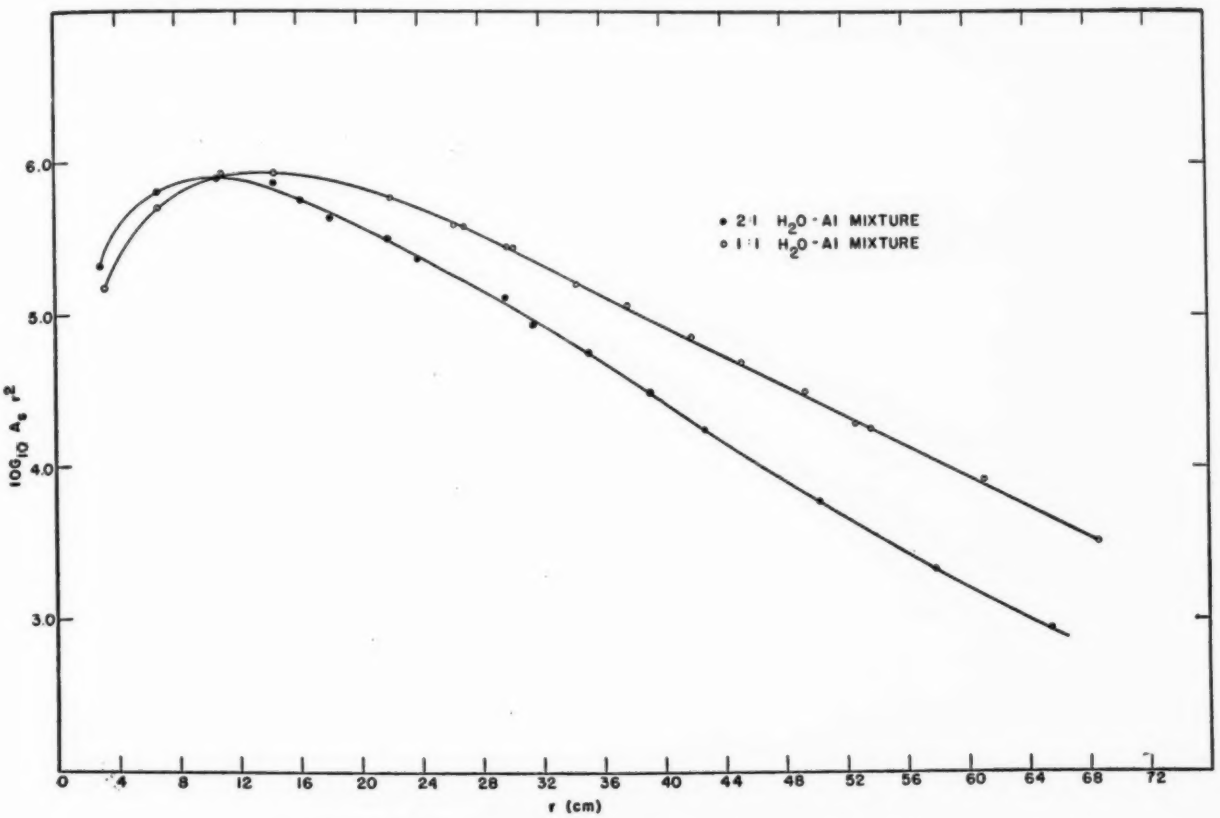


FIG. 2. Plot of the experimental data for the two aluminum water mixtures.

was located 13.3 cm from the boundary. It is assumed that this experimental arrangement gives results corresponding to an infinite cube of the mixture. A discussion of this point is given in paper I.

The foils were counted according to the procedure described in detail in paper I. It should be emphasized, however, that all of the activities here reported correspond to the average value of the activity on the front of the foil with the activity on the back. Also, that geometry corrections have been made for foil size and source size so that the activities A_s quoted are proportional to slowing down densities.

EXPERIMENTAL RESULTS, CALCULATIONS, AND DISCUSSION

Table I gives the experimental data as the function $\log_{10}A_s r^2$ *vs* r, where A_s is the measured saturated activity of the foil (corrected as stated in the foregoing), and r is the distance from the center of the source to the center of the foil. These results are plotted on the graph, Fig. 2, and it is seen that the data fall on reasonably smooth curves. These curves were drawn on quite a large scale and values of $\log_{10}A_s r^2$ read off every two cm. The smoothed values of $A_s r^2$ thus obtained are given in Table II. These experimental functions were extrapolated to infinity using A_s(extrapolated) $= k(e^{-r/\lambda})/r^2$. For the 1:1 mixture the values, $k=6.886\times10^6$ counts/min and $\lambda=8.975$ cm were used, and for the 1:2 mixture we used $k=1.923\times10^6$ counts/min and $\lambda=8.489$ cm. These values were obtained by fitting A_s (extrapolated) to the experimental curves at large r.

TABLE I. Experimental data.

1 volume water:1 volume of aluminum		2 volumes water:1 volume of aluminum	
r(cm)	$\log_{10}A_s r^2$	r(cm)	$\log_{10}A_s r^2$
3.31	5.1830	2.98	5.321
6.85	5.7101	6.78	5.806
10.99	5.9070	10.95	5.906
14.45	5.9269	14.40	5.874
18.67	5.8724	16.31	5.759
22.10	5.7649	18.21	5.640
26.35	5.5929	22.02	5.500
26.89	5.5764	23.93	5.368
29.75	5.4409	29.64	5.114
30.19	5.4348	31.55	4.937
34.39	5.1977	35.13	4.739
37.69	5.0449	39.17	4.493
41.86	4.8383	42.75	4.238
45.19	4.6748	50.37	3.767
49.39	4.4798	57.99	3.320
52.69	4.2786	65.61	2.930
53.72	4.2415		
61.22	3.9047		
68.72	3.5161		
76.22	3.1479		

TABLE II. Interpolated data from $\log_{10} A_s r^2$ curves.

	1 volume aluminum:1 volume water			1 volume aluminum:2 volumes water	
r	$A_s r^2$	$A_s r^4$	r	$A_s r^2$	$A_s r^4$
0	0.0000	0.0000	0	0.0000	0.0000
2	0.6700×10^5	0.00267×10^8	2	0.8800×10^5	0.3520×10^7
4	2.104	0.03373	4	3.3200	0.5312
6	4.179	0.1504	6	3.6664	2.0399
8	6.124	0.3919	8	7.3152	4.6817
10	7.586	0.7586	10	8.0360	8.0360
12	8.356	1.2033	12	7.8782	11.344
14	8.492	1.6644	14	7.1501	14.014
16	8.260	2.1146	16	6.0698	15.538
18	7.709	2.4978	18	5.0414	16.334
20	6.823	2.7292	20	4.0460	16.184
22	5.848	2.8304	22	3.1978	15.477
24	4.920	2.8339	24	2.4687	14.219
26	4.064	2.7472	26	1.8969	12.823
28	3.319	2.6021	28	1.4433	11.315
30	2.685	2.4165	30	1.1025	9.9225
32	2.123	2.1740	32	0.8419	8.6210
34	1.686	1.9490	34	0.6381	7.3764
36	1.340	1.7366	36	0.4838	6.2700
38	1.064	1.5364	38	0.3636	5.2504
40	0.8453	1.3525	40	0.2730	4.3680
42	0.6714	1.1843	42	0.2027	3.5756
44	0.5333	1.0325	44	0.1496	2.9863
46	0.4236	0.8963	46	0.1108	2.3434
48	0.3365	0.7753	48	0.0818	1.8847
50	0.2673	0.6683	50	0.0604	1.5100
52	0.2123	0.5741	52	0.0456	1.2330
54	0.1687	0.4919	54	0.0347	1.0119
56	0.1346	0.4221	56	0.0265	0.8310
58	0.1086	0.3653	58	0.0206	0.6930
60	0.08630	0.3107	60	0.0162	0.5832
62	0.06902	0.2653	62	0.0128	0.4920
64	0.05521	0.2261	64	0.0102	0.4178
66	0.04416	0.1924	66	0.0080	0.3485
68	0.03540	0.1637			
70	0.02831	0.1387			
72	0.02265	0.1174			
74	0.01811	0.09917			
76	0.01452	0.08387			
78	0.01162	0.07069			
80	0.009290	0.05945			

The value of $\bar{r}^2$, the quantity which we sought to obtain, is given by the integral

$$\bar{r}^2 = \frac{\int_0^\infty A_s r^4 dr}{\int_0^\infty A_s r^2 dr}.$$

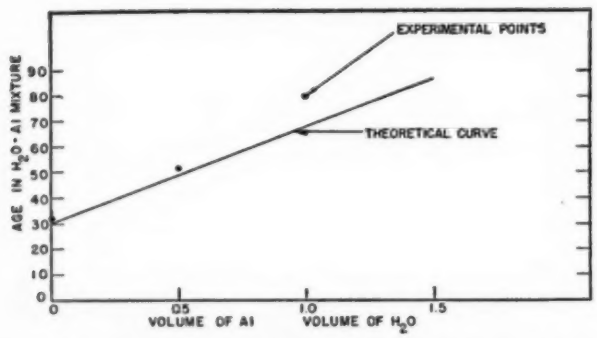


FIG. 3. Comparison of a theoretical function for the age with the experimental values.

TABLE III.

One volume of Al to one volume H_2O
measured $\int_0^{80} A_s r^2 dr = 183.1\times10^5$
extrapolated $\int_{80}^\infty A_s r^2 dr = 0.08\times10^5$
measured $\int_0^{80} A_s r^4 dr = 83.68\times10^8$
extrapolated $\int_{80}^\infty A_s r^4 dr = 0.67\times10^8$
$\bar{r}^2 = 460.7$ cm^2
$\tau = \bar{r}^2/6 = 76.8$ cm^2
One volume of Al to two volumes H_2O
measured $\int_0^{66} A_s r^2 dr = 137.2\times10^5$
extrapolated $\int_{66}^\infty A_s r^2 dr = 0.07\times10^5$
measured $\int_0^{66} A_s r^4 dr = 40.44\times10^8$
extrapolated $\int_{66}^\infty A_s r^4 dr = 0.39\times10^8$
$\bar{r}^2 = 297.4$ cm^2
$\tau = \bar{r}^2/6 = 49.6$ cm^2

The following Table III gives our values for the function, $\bar{r}^2$ to indium resonance.

It is seen that only 1% of our $\bar{r}^2$ values is due to extrapolated area. From a consideration of total count and of the measurement of r at the experimental points, it would seem that the values of $\bar{r}^2$ and τ given in Table III are good to the order of 1%. Here τ is the neutron age defined as $\bar{r}^2/6$.

Table IV gives our age values from fission energies to indium resonance enhanced by an estimated age increment from indium resonance energies to thermal energies.

The graph, Fig. 3, gives a theoretical function[1] for τ *vs* the ratio vol. Al/vol. H_2O together with our experimental points. It is seen that the experimental point for pure water falls close to the theoretical function, but that the values of the age τ for the two aluminum

TABLE IV.

Vol. Al/ vol. H_2O	Experimental age to indium resonance	Theoretical age indium resonance to thermal	Total age to thermal energies
0	30.8 cm^2	1 cm^2	31.8 cm^2
0.5	49.6 cm^2	2 cm^2	51.6 cm^2
1.0	76.8 cm^2	3 cm^2	79.8 cm^2

[1] Given in the Oak Ridge National Laboratory Report Mon P-219 by Weinberg, Soodak, Dismuke, and Arnette.

water mixtures fall progressively further from this curve. Even though fair agreement with experiment is obtained for pure water, the theoretical calculation procedure is not really adequate. This is brought out by the fact that various attempts to improve the calculation for pure water have led to poorer rather than better agreement.[2] It is thus of interest to note the progressively increasing disagreement between experiment and this first-order theory for aluminum water mixtures.

[2] This calculation has been made recently at a number of laboratories; see, for example, J. Certaine and R. Aronson, Rept. NDA 15C-40.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Rectification Properties of Metal Semiconductor Contacts

E. H. Borneman, R. F. Schwarz, and J. J. Stickler
Philco Corporation, Philadelphia, Pennsylvania

(Received December 22, 1954; revised manuscript received March 17, 1955)

Metal semiconductor contacts of a number of different metals were made on *n*- and *p*-type germanium using jet etching and plating techniques. Current voltage curves taken on 12 of these metals on 5 ohm-cm *n*-type germanium showed rectification which follows the diode equation $J=J_0(e^{qV/kT}-1)$. No correlation was found between the reverse saturation current densities of these diodes and such properties of the metals as work function, electromotive force, etc. For those metal contacts possessing the lowest saturation current densities, calculations indicated the current crossing the contact was to a large percent hole current and that the magnitude of the hole current was controlled primarily by the geometry of the diode. All metals plated on 5 ohm-cm *p*-type germanium produced ohmic contacts of resistivity comparable to the spreading resistance expected for the diode geometry used.

For indium diodes, a study of rectification *versus* resistivity indicated that the barrier produced on both *n*- and *p*-type germanium with plated contacts is one to electron flow rather than hole flow. When the assumption of only hole current crossing the barrier was made, it was shown that the I–V curves calculated from the diode theory, for different resistivities of germanium, were in qualitative agreement with the measured curves. Curves of zero voltage conductance *versus* temperature for different resistivities of germanium were also found to be in good agreement with those calculated on the assumption of all hole current.

INTRODUCTION

As reported previously by Bradley *et al.*,[1] when indium is electroplated onto freshly etched *n*-type germanium, rectifying contacts are obtained. Used as diodes, or as emitters and collectors in transistors, the electrical behavior of these contacts is very similar to indium fused *p-n* junctions even though no diffusion of impurities has taken place. This discovery instigated the following study, a survey of the rectifying properties of various metals electroplated on both *n*- and *p*-type germanium.

The rectification of the indium contacts has been attributed to a surface barrier to electrons (see Fig. 1). For such a model, theory predicts that the electron current at the contact I_n should be of the form

$$I_n=I_{ns}(e^{qV/kT}-1), \tag{1}$$

where V is the applied voltage, q is the value of the electronic charge, k is Boltzmann's constant, and T is the absolute temperature. The electronic saturation current, I_{ns}, should be of the form

$$I_{ns}=AT^2e^{-q\Phi_0/kT}, \tag{2}$$

where A is a constant and Φ_0 is the potential difference in electron volts between the Fermi level and the bottom of the conduction band at the germanium-metal interface. Thus, Φ_0 is a measure of the height of the surface barrier. The hole current at the metal contact, I_p, should also be of the form

$$I_p=I_{ps}(e^{qV/kT}-1), \tag{3}$$

where I_{ps}, the hole saturation current, should be limited by the ability of the holes to diffuse in the bulk of the germanium. Since the hole current is limited by this diffusion process in both the *p-n* junction and the surface barrier model, the value of I_{ps} can be calculated in the same way for both. Its actual value depends on the recombination rate of the holes in the bulk of the germanium and at the surface, as well as on the geometry of the contact. For a circular contact of radius a on a semi-infinite slab of *n*-type germanium, the hole saturation current would be

$$I_{ps}=4kT\mu_p a p_{eq}, \tag{4}$$

if the bulk and surface recombination are negligible. Here μ_p is the mobility of the holes, and p_{eq} is the equilibrium value of the hole concentration in the bulk.

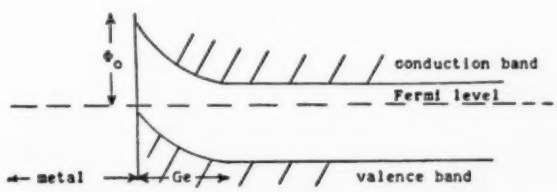


FIG. 1. Diagram of the electron barrier at germanium-metal interface.

[1] W. E. Bradley, Proc. Inst. Radio Engrs. 41, 1702 (1953).

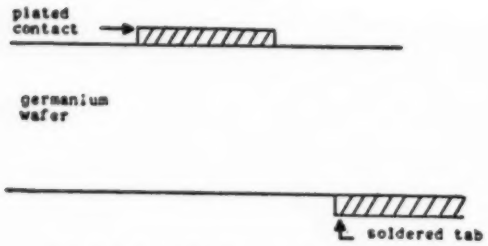


FIG. 2. Schematic diagram of the geometry of the contacts studied.

This geometry approximates that of the contacts studied (see Fig. 2).[2]

The total current, I, is given by the sum of I_n and I_p and should therefore also be of the form

$$I = I_s(e^{qV/kT} - 1), \tag{5}$$

where I_s, the total saturation current, is the sum of I_{ns} and I_{ps}. Thus Eq. (4) represents the minimum value of the saturation current which one could expect to obtain for these diodes on n-type material. Recombination of the holes would make I_{ps} larger than this minimum value.

EXPERIMENTAL MATERIALS AND TECHNIQUES

The metal semiconductor contacts were formed by an application of techniques described in detail by Tiley and Williams.[3] In brief, the method consists of directing a jet of an aqueous solution of a metal salt through a glass nozzle onto a germanium wafer, held perpendicular to the direction of the stream. A stainless steel holder supporting the germanium blank permits electrical contact to be made to the wafer. The electrical circuit is completed by a platinum wire which serves as the second electrode in the salt solution. In general, a 0.1 N metal salt solution was used, with another suitable electrolyte added for the purpose of increasing the solution conductivity.

In order to insure the best possible contact of the metal to the germanium, the germanium was electrolytically etched by making it the anode in the cell, the platinum wire then serving as the cathode. After suitably etching the germanium, plating was begun by reversing the potential, thus making the germanium the cathode. Metal ions would then deposit on the same region which was etched, forming a metal semiconductor contact. The plating thickness could be regulated by varying either the current density or the time. After plating, a chemical etch was used to clean the sample, and to give a sharply defined region of metal contact.

In practice, good plating was obtained with a solution conductivity of about 0.05 $ohm^{-1}-cm^{-1}$. Attempts to plate a large number of different metals were made with varying degrees of success. Table I lists some of the metals as well as the composition of the solutions. Some difficulty was encountered with those elements having electrochemical potentials much below hydrogen; namely, platinum, copper, silver, and gold, due to the tendency of these metals to spontaneously deposit upon contact with the germanium. However, by adjusting the plating conditions it was possible to obtain deposits of these metals having good electrical proper-

TABLE I. Composition of solutions used to plate various metals on germanium.

Plated metal	Solution components	Grams/liter
Indium	$In_2(SO_4)_3$	8.6
	H_2SO_4	2.45
Zinc	$ZnSO_4 \cdot 7H_2O$	50
	NH_4Cl	5
	$NaC_2H_3O_2$	5
Cadmium	$CdCl_2$	4.5
	NH_4Cl	5
Tin	$SnCl_2$	6.4
	HCl	6.1
Lead	$Pb(OH)_2 \cdot 2PbCO_3$	150
	HF	120
	H_3BO_3	105
Copper	$CuSO_4$	8
	H_2SO_4	4.9
Silver	$AgNO_3$	0.34
	$NaNO_3$	17
Gold	$AuCl_3$	10
	NaCl	2.9
Rhodium	$RhCl_3 \cdot 4H_2O$	8.5
	H_2SO_4	4.9
Nickel	$NiCl_2 \cdot 6H_2O$	11.8
	HCl	2.5
Platinum	$Pt(NH_3)_2(NO_2)_2$	5
	NH_4NO_3	50
	$NaNO_2$	5
Chromium	CrO_3	16.6
	Na_2CO_4	0.24
Tungsten	WO_3	32
	Na_2CO_3	48.5
	$NiCO_3$	0.09
Arsenic	As_2O_3	3.3
	Na_2CO_3	5.3
Antimony	SbF_3	6
	HF	4
Bismuth	BiOCl	20
	HCl	73.2

[2] (a) Equation (4) is calculated in the same manner that one would calculate the spreading resistance of such a contact. In fact, in the absence of recombination, I_{ps} should be inversely proportional to the resistance of an ohmic contact with the same geometry. For typical dimensions of the diodes used, this hole current due to geometrical effects is larger and therefore more important than that due to recombination calculated by Shockley for p–n junctions; (b) W. Shockley, Bell System Tech. J. **28**, 435 (1949). This implies that the previous estimate calculated by Schwarz and Walsh for hole current in the indium contacts is too low; (c) R. F. Schwarz, and J. F. Walsh, Proc. Inst. Radio Engrs. **41**, 1715 (1953).

[3] J. W. Tiley and R. A. Williams, Proc. Inst. Radio Engrs. **41**, 1706 (1953).

ties. The physical nature of the plated deposit depended on the plating current density as well as on the composition of the solution. Using rhodium as an example, it was possible to produce either a thin bright deposit, or a soft granular deposit depending upon the magnitude of the current density. Difficulty was experienced in obtaining good deposits of chromium and tungsten, which was attributed to excess oxidation of the deposition region.

The germanium used in all of these experiments was single crystal, 110 orientated, and of approximately 100 μsec lifetime.

I-V CHARACTERISTICS ON 5 ohm-cm *n*-TYPE MATERIAL

All metals plated onto 5 ohm-cm *n*-type germanium made good rectifying contacts. Below 0.1 volt in either direction the current voltage curves followed closely the relation Eq. (5). The saturation current, I_s, was found to depend on the physical nature of the deposited metal, bright adhering deposits giving the lowest values. If

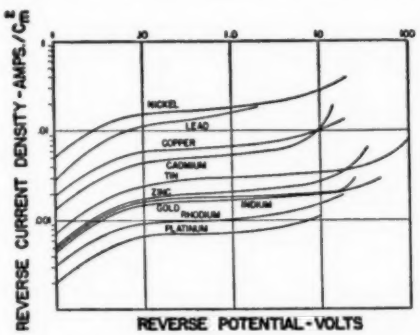


FIG. 3. I–V curves in the reverse direction of metals plated on 5 ohm-cm *n*-type germanium.

the plating conditions were kept the same, I_s would not vary by more than a factor of two or three from sample to sample.

Reverse characteristics for twelve metals are shown in Fig. 3. For each metal the curve is representative of the plating conditions which gave the lowest obtained value of the saturation current per unit area of contact. The area of electrical contact was determined from measurements of capacitance *vs* voltage.[2(e)] For the diodes reported on, this area was between 10^{-3} and 2×10^{-3} cm², and agreed closely with the visual area of contact. Equation (4) would thus predict that the minimum saturation current which could be obtained for these diodes would be in the neighborhood of 9×10^{-4} to 6×10^{-4} amp/cm². This value may be compared with the minimum value of 7×10^{-4} amp/cm² obtained for the platinum diodes. It should be stressed that Eq. (4) pertains only to the hole saturation current. Since this equation predicts a current about equal to the measured total saturation current, then for these diodes it appears that the barrier is sufficiently high to make the electron current negligible compared to the hole current.

TABLE II. Comparison of saturation current density of several metal-germanium diodes to the work function of the metal.

Metal	Saturation current amp/cm²	Metal W.F.[a]
Platinum	0.00049	5.2
Rhodium	0.00095	4.6
Gold	0.0012	4.8
Indium	0.0014	4.0
Zinc	0.0016	4.3
Tin	0.0023	4.5
Cadmium	0.0038	4.1
Copper	0.0092	5.2
Lead	0.011	4.2
Nickel	0.065	5.0

[a] Herbert B. Michaelson, J. Appl. Phys. 21, 536 (1950).

Also, recombination does not seem to be playing an important role. For rhodium, gold, indium, and zinc the values of I_s all fall within a factor of two of the limits set for the minimum value of I_{ps}. Since for the same plating conditions the values of I_s cannot be repeated within less than a factor of two, the current in these diodes may also be primarily hole current. The excess current in these particular samples may be due to recombination of holes at the surface or the slight differences in geometry from that implied by Eq. (4). On the other hand, for the metals with the highest saturation currents, i.e., nickel, lead, and copper, it seems most likely that some electron current is present, thus indicating a lower surface barrier.

Some attempts were made to correlate this data with different properties of the metals. Referring to Table II, one sees that there does not appear to be any obvious correlation between the metal work function and the measured saturation current of the diodes. Other properties of the metals, such as electrochemical potential, melting temperature, doping properties, etc.,

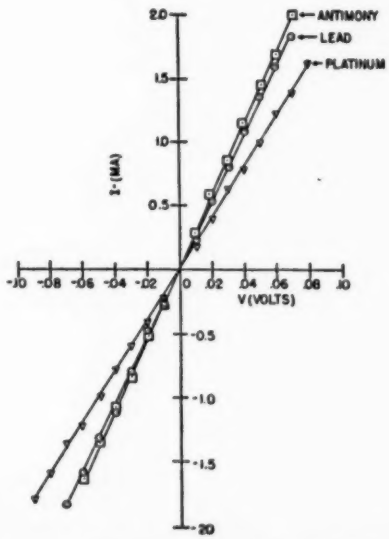


FIG. 4. I–V curves of metals plated on 5 ohm-cm *p*-type germanium.

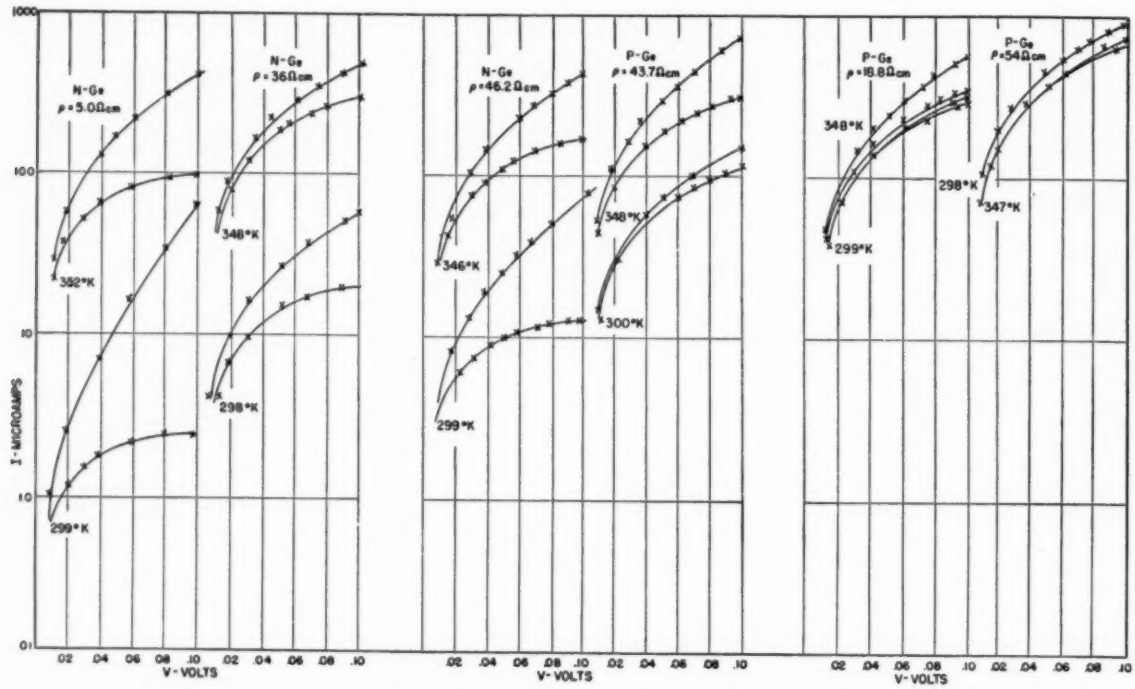


FIG. 5. I–V characteristics of indium-germanium diodes for various resistivities and temperatures.

also show no obvious correlation. In general, it seems that the saturation current is affected more by the physical nature of the deposit than by the metal itself.

I-V CHARACTERISTICS ON 5 ohm-cm *p*-TYPE GERMANIUM

No rectification was obtained on 5 ohm-cm *p*-type germanium with any metal tried. Figure 4 shows typical I–V curves at low voltages for three of the metal contacts. The linearity of the I–V relation extended to higher voltages than shown on the graph. Table III lists the average dc resistance measured for the different metal contacts. These resistances may be compared with the spreading resistance of a circular contact on a semi-infinite slab of germanium. For these contacts this spreading resistance should be between 50 and 70 ohms. Most of the metals show resistances somewhat lower than 50 ohms. This low resistance is probably due to the proximity of the soldered tab to the metal electrode, and indicates that I_{ps} calculated from Eq. (4) for *n*-type samples is probably somewhat too small.

TABLE III. Resistance of metal contacts plated on 5 ohm-cm *p* germanium.[a]

Metal	*R* in ohms
Tin	70
Antimony	40
Arsenic	100
Copper	40
Cadmium	40
Rhodium	30
Lead	30
Bismuth	30
Platinum	30
Gold	40
Indium	30
Zinc	40
Nickel	40

[a] Diameters of metal dots were all approximately 0.015 in.

I-V CHARACTERISTICS OF THE INDIUM CONTACTS *VS* RESISTIVITY AND TEMPERATURE

A study was made of the indium contacts on various resistivities of the germanium covering the range from 5 ohm-cm *n*–type germanium through intrinsic to 5 ohm-cm *p* type. Current-voltage curves on each resistivity sample were taken for various temperatures. In Fig. 5, I–V curves are shown for each resistivity at two different temperatures.

From Fig. 5 it can be seen that as one goes from low resistivity *n* type through intrinsic to low resistivity *p* type (all at room temperature), the rectification ratio steadily decreases. On 18 ohm-cm *p* type some rectification still persists, while on 5 ohm-cm *p* type no rectification is detected. It should be emphasized that the rectification is always of the same polarity, the forward direction corresponding to the germanium being negative. As the temperature is increased, the rectification ratio decreases on *n*-type samples and increases on *p*-type samples.

For comparison purposes, a similar study was made on indium fused *p-n* junctions. In the range of resistivities and temperatures studied, essentially identical results were obtained for the indium plated and indium fused contacts. Figure 6 shows curves of the saturation current *versus* the reciprocal of the temperature for a plated contact and a *p-n* junction, both made on 5 ohm-

cm n-type germanium. The logarithm of the saturation current *versus* the reciprocal of the temperature are straight lines with the same slope for both types of contacts. For a p-n junction, the slope of this curve should be given approximately by $(-q\Phi_G/k)$ where Φ_G is the gap width of the germanium. The slope of the p-n junction data gives a value for Φ_G of 0.67 volt. If the current in the plated contact were primarily hole current, one would expect to obtain the same value for this slope. On the other hand, if the current were primarily electron current, one would expect from Eq. (2) that the slope would be given approximately by $(-q\Phi_0/k)$. The fact that the same slope is obtained for the plated contacts as for the p-n junctions indicates either that the current is carried primarily by holes, or that the barrier height, Φ_0, is of the order of 0.67 volt. In the latter event, the barrier would be sufficiently high that the current would of necessity be primarily hole current, for the 5 ohm-cm sample.

THEORETICAL CONSIDERATIONS

The data presented strongly suggests that for all metals plated on both n- and p-type germanium there exists a potential barrier to electron flow from the semiconductor to the metal, but none for hole flow. For many of the metals plated under optimum conditions, it appears that this barrier may be sufficiently high to block the electron current almost entirely. This viewpoint is substantiated for the indium contacts by the fact that, at least qualitatively, it leads to an adequate description of the resistivity and temperature data.

This may be seen to be the case, if the following assumptions are made.

(a) Only valence electrons can cross the metal contact (i.e., the current is 100% hole current at the contact).

(b) The lifetime of the holes is infinite. This implies that there is only hole current everywhere in the bulk. (Actually recombination would cause more and more of the hole current to transfer to electron current as one proceeds through the bulk.)

(c) In the bulk there is no space charge. Therefore,

$$n-n_{eq}=p-p_{eq}, \tag{6}$$

where n and p are the electron and hole densities, respectively, and n_{eq} and p_{eq} are their equilibrium values.

(d) The hole density at the soldered tab is p_{eq}. (This is probably only an aproximation of the actual facts, and implies that no voltage drop occurs across this contact.)

(e) In the bulk, the electron and hole current densities, j_p and j_n, are given by

$$j_p=q\mu_p pE-kT\mu_p\nabla p, \tag{7}$$

$$j_n=q\mu_n nE+kT\mu_n\nabla n, \tag{8}$$

where μ_p and μ_n are the mobilities of holes and electrons, respectively, and E is the electric field.

In appendix A the following equation is derived from the above assumptions.

$$I_p=CkT\mu_p\left[(n_{eq}+p_{eq})\left\{\left(1+\frac{4p_{eq}n_{eq}}{(n_{eq}+p_{eq})^2}(e^{qV/kT}-1)\right)^{\frac{1}{2}}-1\right\}\right.$$
$$\left.-(n_{eq}-p_{eq})\ln\left\{1+\frac{(n_{eq}+p_{eq})}{2n_{eq}}\left[\left(1+\frac{4p_{eq}n_{eq}}{(n_{eq}+p_{eq})^2}(e^{qV/kT}-1)\right)^{\frac{1}{2}}-1\right]\right\}\right]. \tag{9}$$

Here I_p is the total current crossing the contact, and C is a constant factor which depends only on the geometry of the diode.

If V is not too large, this equation may be shown to reduce to the following equations for the limiting cases stated.

(a) $n_{eq}/p_{eq}\gg 1$ (i.e., n-type material)

$$I_p=CkT\mu_p p_{eq}(e^{qV/kT}-1) \tag{10a}$$

(b) $n_{eq}/p_{eq}=1$ (i.e., intrinsic material)

$$I_p=2CkT\mu_p p_{eq}(e^{qV/2kT}-1), \tag{10b}$$

(c) $n_{eq}/p_{eq}\ll 1$ (i.e., p-type material)

$$I_p=Cq\mu_p p_{eq}V. \tag{10c}$$

Equation (10a) is that which would be obtained if one assumed that the hole current resulted only from the diffusion term of Eq. (7). On the other hand, (10c) is that which would be obtained if one assumed that the hole current resulted only from the field term of Eq. (7).

For low voltages, Eq. (9) thus predicts that the I–V curves should vary in shape from that of an ideal rectifier with an exponent of q/kT, to purely ohmic curves as the resistivity goes from low n type through intrinsic to low p type. Equation (9) is plotted for various resistivities in Fig. 7. The current densities are shown to scale within an arbitrary factor. These curves may be compared with the room temperature experimental curves of Fig. 5. Qualitatively and quantitatively, the curves behave very much as predicted by Eq. (9), except for the following discrepancies:

(a) The rectification ratio is usually somewhat less than predicted by Eq. (9).

(b) Relative to the currents in the p-type samples, the currents in the n-type samples are slightly larger than the theory predicts.

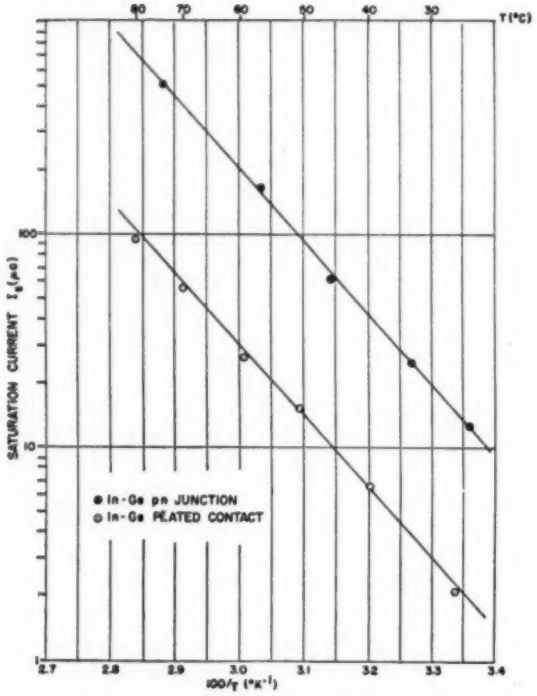


FIG. 6. Saturation current as a function of temperature for a p-n junction and a plated contact.

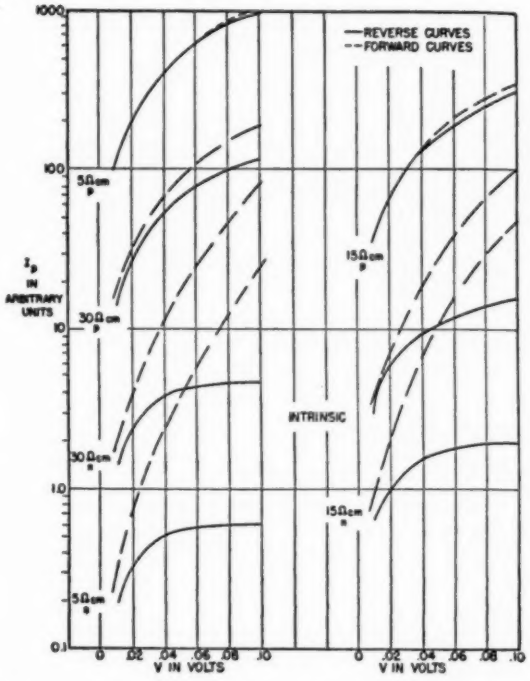


FIG. 7. Theoretical I–V curves for various resistivities of germanium assuming all hole current, infinite lifetime and $T = 300°K$.

If recombination were included in the theory, the theoretical curves (Fig. 7) would change in such a manner as to reduce the above discrepancies.

From Eqs. (10a), (10b), and (10c) it can be seen that the zero voltage conductance is given by

$$dI_p/dV \mid_{V=0} = Cq\mu_p p_{eq} \tag{11}$$

for all resistivities. Thus the zero voltage conductance should vary as $\mu_p\, p_{eq}$. Figure 8 shows theoretical curves of $\mu_p\, p_{eq}$ *versus* temperature for various resistivities. These curves compare favorably with the experimental curves of the zero voltage conductance *versus* temperature in Fig. 9.

CONCLUSIONS

All metals that were plated on 5 ohm-cm n-type germanium gave good rectification. No obvious correlation was found between the reverse saturation currents of these diodes and such properties of the metals as work function, electromotive force, etc. Calculations indicate that those metal diodes which give the lowest value of saturation current density (Pt, Au, Rh, In, and Zn) have a large percentage of hole current, the magnitude of the hole current being controlled primarily by the geometry of the diode.

All metals plated on 5 ohm-cm p-type germanium gave ohmic contacts of resistance comparable to the spreading resistance which one would expect for the geometry used.

For the indium diodes, the variation of the current voltage curves with resistivity of the germanium showed rectification to persist from n–type germanium through intrinsic to p-type germanium. The rectification was always of the same polarity, but the rectification ratio steadily decreased as the germanium was made more

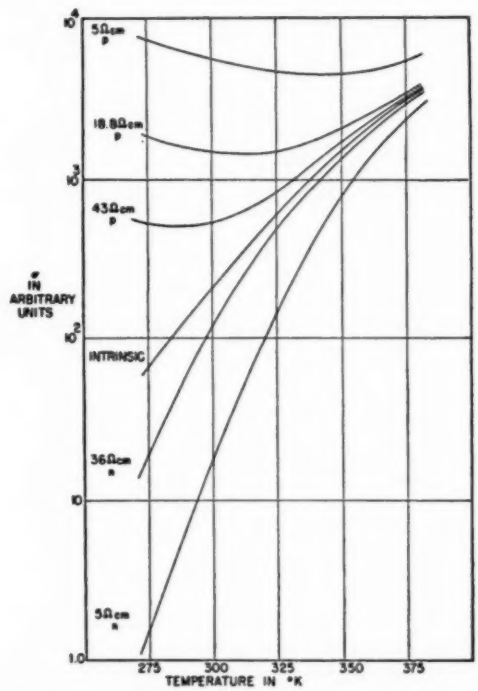


FIG. 8. Theoretical curves of zero-voltage conductance *versus* temperature for germanium assuming all hole current and infinite lifetime.

p type. An analysis of these curves substantiates the conclusion that the current across the plated contact was carried primarily by holes. Curves of the zero voltage conductance *versus* temperature for the various resistivities also agree qualitatively with those calculated from theory based on the assumption of all hole current. It thus appears that it is relatively easy to establish a very high barrier to electron flow on germanium surfaces. On the other hand, it has not been possible to obtain a barrier to hole flow with any metal tried.

Common models of the surface barrier, such as surface states or work function differences, would indicate that the electron barrier should stay constant or slightly increase as the resistivity is increased from low n type toward intrinsic. For a given metal, an upper limit on the amount of electron current should therefore be given by the lowest resistivity curve measured. On the indium contacts, the lowest resistivity curve measured was for n-type 0.1 ohm-cm material. With this material the current voltage curve saturated at about 0.2 μa. Therefore, for the indium diodes of higher resistivity than 0.1 ohm-cm and with an area of contact of about 10^{-3} cm^2, the electron current in the reverse direction at room temperature should always be less than this number.

ACKNOWLEDGMENTS

The writers are indebted to Miss Betty Knoeri for making many of the measurements reported in this paper, and to Mr. Joseph Walsh for the data presented in Fig. 6.

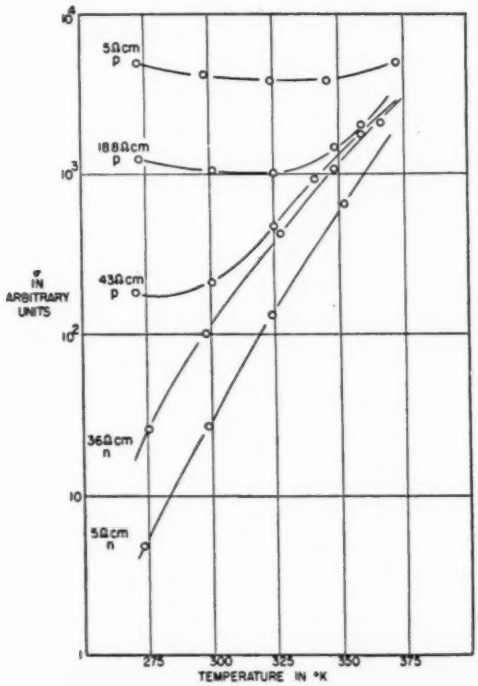


FIG. 9. Experimental curves of zero-voltage conductance *versus* temperature for germanium.

APPENDIX A. MATHEMATICAL DERIVATION OF EQ. (9)

The derivation of Eq. (9) is based on the assumptions given on page 1023.

$$n-n_{eq}=p-p_{eq} \tag{A1}$$

$$j_n=q\mu_n nE+kT\mu_n\nabla n \tag{A2}$$

$$j_p=q\mu_p pE-kT\mu_p\nabla p. \tag{A3}$$

Since j_n is assumed to be everywhere zero, (A2) may be solved for E in the bulk.

$$E=\frac{-kT}{q}\frac{\nabla n}{n}. \tag{A4}$$

Upon integrating from the soldered tab to the metal contact, one obtains the voltage V_1 appearing across the bulk.

$$V_1=\frac{kT}{q}\ln\frac{n_0}{n_{eq}}=\frac{kT}{q}\ln\left(1+\frac{p_0-p_{eq}}{n_{eq}}\right), \tag{A5}$$

where p_0 and n_0 are the densities of holes and electrons, respectively, just outside the space-charge region at the metal contact. The voltage appearing across the barrier is $V-V_1$, and this is related to p_0 by

$$V-V_1=\frac{kT}{q}\ln\frac{p_0}{p_{eq}}. \tag{A6}$$

Combining (A5) and (A6),

$$V=\frac{kT}{q}\ln\left[\frac{p_0}{p_{eq}}\left(1+\frac{p_0-p_{eq}}{n_{eq}}\right)\right] \tag{A7}$$

$$=\frac{kT}{q}\ln\left[1+\frac{\delta p}{p_{eq}}+\frac{\delta p}{n_{eq}}+\frac{\delta p^2}{n_{eq}p_{eq}}\right],$$

where δp represents (p_0-p_{eq}). On solving for δp, one obtains

$$\delta p=\frac{n_{eq}+p_{eq}}{2}\times\left(-1+\left[1+\frac{4p_{eq}n_{eq}}{(n_{eq}+p_{eq})^2}(e_q{}^{V/kT}-1)\right]^{\frac{1}{2}}\right). \tag{A8}$$

Returning now to the hole current, we have from (A1), (A3), and (A4).

$$j_p=-kT\mu_p\left(\frac{p}{n}+1\right)\nabla p$$

$$=\frac{-kT\mu_p(2p+n_{eq}-p_{eq})\nabla p}{(p+n_{eq}-p_{eq})}. \tag{A9}$$

Since the electron current is presumed to be every-

where zero, continuity of current requires that

$$\nabla\cdot j_p=0=-kT\mu_p \times\left[\frac{(2p+n_{eq}-p_{eq})\nabla^2 p}{(p+n_{eq}-p_{eq})}+\frac{(n_{eq}-p_{eq})\nabla p\cdot\nabla p}{(p+n_{eq}-p_{eq})^2}\right]. \quad \text{(A10)}$$

Thus,

$$\nabla^2 p=-\frac{(n_{eq}-p_{eq})\nabla p\cdot\nabla p}{(p+n_{eq}-p_{eq})(2p+n_{eq}-p_{eq})}.$$

In Appendix B, a proof is given that Eqs. (A9) and (A10) always lead to the following result for the total current I_p crossing the plated contact:

$$I_p=CkT\mu_p\left[2\delta p-(n_{eq}-p_{eq})\ln\left(1+\frac{\delta p}{n_{eq}}\right)\right]. \quad \text{(A11)}$$

C is a constant depending only on the geometry. From (A8) and (A11) we obtain

$$I_p=CkT\mu_p\left[(n_{eq}+p_{eq})\left\{\left(1+\frac{4p_{eq}n_{eq}}{(n_{eq}+p_{eq})^2}(e^{qV/kT}-1)\right)^{\frac{1}{2}}-1\right\} -(n_{eq}-p_{eq})\ln\left\{1+\frac{(n_{eq}+p_{eq})}{2n_{eq}}\left[\left(1+\frac{4p_{eq}n_{eq}}{(n_{eq}+p_{eq})^2}(e^{qV/kT}-1)\right)^{\frac{1}{2}}-1\right]\right\}\right].$$

APPENDIX B. DERIVATION OF EQ. (A11) FROM EQS. (A9) AND (A10)

Consider generalized orthogonal coordinates where the three elements of length are h_1dq_1, h_2dq_2, and h_3dq_3. Let dq_1 be along a normal to surfaces of constant hole density. Then Eq. (A10) becomes explicitly

$$\frac{1}{h_1h_2h_3}\frac{d}{dq_1}\left[\frac{h_2h_3}{h_1}\frac{dp}{dq_1}\right] =\frac{(n_{eq}-p_{eq})(dp/dq_1)^2}{h_1^2(p+n_{eq}-p_{eq})(2p+n_{eq}-p_{eq})}. \quad \text{(B1)}$$

The solution to Eq. (B1) is

$$2p-(n_{eq}-p_{eq})\ln(p+n_{eq}-p_{eq}) =C_1\int F(q_1)dq_1+C_2, \quad \text{(B2)}$$

where C_1 and C_2 are constants of integration, and $F(q_1)$ is given by the dependence of h_1/h_2h_3 on q_1.
Let

$$\int F(q_1)dq_1=f(q_1). \quad \text{(B3)}$$

Upon evaluating the constants at the two boundaries, Eq. (B2) becomes

$$\frac{2(p-p_{eq})-(n_{eq}-p_{eq})\ln[1+(p-p_{eq})/n_{eq}]}{2(p_0-p_{eq})-(n_{eq}-p_{eq})\ln[1+(p_0-p_{eq})/n_{eq}]}=\frac{f(q_1)-f(q_{1eq})}{f(q_{10})-f(q_{1eq})}. \quad \text{(B4)}$$

Here $f(q_{10})$ and $f(q_{1eq})$ are the values of $f(q_1)$ evaluated at the surfaces for which $p=p_0$ and $p=p_{eq}$, respectively. From (A9) and (B4),

$$j_p=-kT\mu_p\{2\delta p-[(n_{eq}-p_{eq})\ln(1+\delta p/n_{eq})]\} \times\frac{df(q_1)/dq_1}{h_1[f(q_{10})-f(q_{1eq})]}. \quad \text{(B5)}$$

Thus, upon integrating (B5) over the area of the plated contact,

$$I_p=kT\mu_p\left[2\delta p-(n_{eq}-p_{eq})\ln\left(1+\frac{\delta p}{n_{eq}}\right)\right] \left\{\frac{-\int\int_{\text{plated contact}}\frac{1}{h_1}\frac{df(q_1)}{dq_1}dq_2dq_3}{f(q_{10})-f(q_{1eq})}\right\},$$

where the expression in the right-hand bracket equals C.

Plasma Oscillations in Electron Beams*

K. G. HERNQVIST
RCA Laboratories, Princeton, New Jersey
(Received December 24, 1954)

Self-sustained ion and electron plasma oscillations have been observed simultaneously in an ion neutralized electron beam traveling in a transverse magnetic field. The observed frequencies of oscillation agree with the Langmuir-Tonks theory for plasma oscillations. Electron plasma oscillations occur when the electron plasma frequency equals the gyro frequency of a circular orbit in the magnetic field.

TWO characteristic plasma frequencies can be theoretically derived for an ion-neutralized electron beam.[1] These are the ion plasma frequency

$$f_p = 0.679 \times 10^6 g \left(\frac{j}{M(V)^{\frac{1}{2}}} \right)^{\frac{1}{2}} \quad (1)$$

and the electron plasma frequency

$$f_e = 29.3 \times 10^6 g \left(\frac{j}{(V)^{\frac{1}{2}}} \right)^{\frac{1}{2}}, \quad (2)$$

where g is a geometrical factor, j is the beam current density, V is the beam voltage, and M is the molecular weight of the ions. MKS-units are used throughout unless otherwise specified. Self-sustained ion plasma oscillations have previously been studied in ion-neutralized electron beams.[1,2] The concept of electron plasma frequency of an electron beam is frequently used in traveling wave tube theory. However, to the author's knowledge, self-sustained oscillations at this frequency in an electron beam under high vacuum conditions have never been reported. In an experiment done at these laboratories, self-sustained oscillations of the two plasma frequencies have been observed simultaneously.

Figure 1 shows the actual tube structure. An electromagnet having two conically-shaped pole pieces, touching at their tips, gives a magnetic field B essentially following the relation $B = \text{const} \times r^{-n}$ where r is the radial distance from the cone-tips and n a constant slightly less than 1. A rectilinear flow, gridded electron gun injects an electron beam tangentially into the electrically field-free toroidal-shaped region limited by the outer shielding ring S, the cone tips T, and the end-plates P. All tube parts are made of nonmagnetic material except the pole pieces. The tube is continuously pumped, the pressure being held below 5×10^{-6} mm Hg.

For zero magnetic field, most of the beam current is collected by S. As the magnetic field is increased, a point is finally reached where the electron beam is caused to spiral slowly in towards the cone tips and a transition of beam current from S to T takes place. The current transition is quite sharp, taking place within about 10% magnetic field strength variation. During this current transition, high-frequency oscillations between 300 and 500 mc could be picked up by a receiver capacitively coupled to any of the electrodes surrounding the beam region. Also, oscillations of much lower frequency (between 1.5 and 3 mc) were observed simultaneously with the high-frequency oscillations by means of a second receiver coupled to the same electrode. In Fig. 2 these two frequencies are

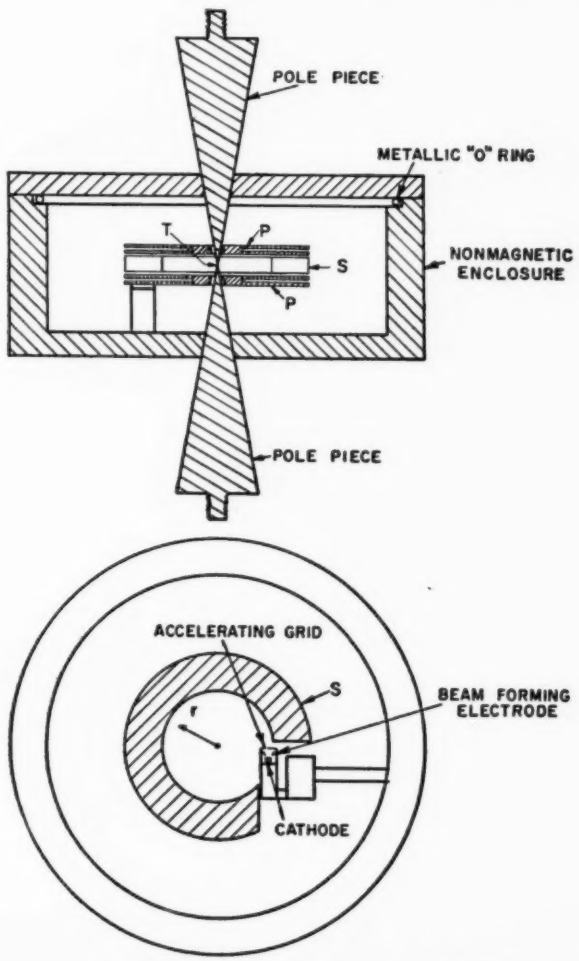


FIG. 1. Scale drawing of essential tube parts.

* Work supported by the Signal Corps of the U. S. Army under contract.

[1] K. G. Hernqvist, J. Appl. Phys. **26**, 544 (1955).

[2] E. G. Linder and K. G. Hernqvist, J. Appl. Phys. **21**, 1088 (1950).

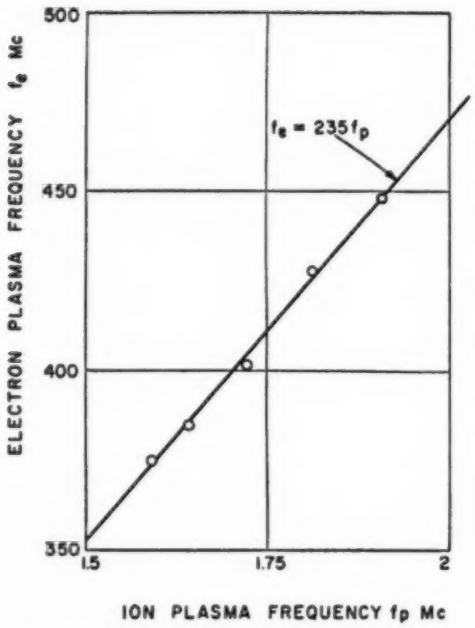


FIG. 2. Relation between the two frequencies of oscillation observed simultaneously.

plotted against each other. Each point is taken at a different value of beam voltage and beam current, the magnetic field being adjusted for maximum oscillation amplitude.

From Eqs. (1) and (2) using an average value of $M=30$ for air, one obtains

$$f_e=235 f_p, \tag{3}$$

the factor 235 being the square root of the ion-to-electron mass ratio. It is seen from Fig. 2 that the observed points follow this relation quite closely. Neither the actual current density of the spiraling beam nor the value of the geometry factor g are known ($g=1$ for longitudinal oscillations and $g=1/\sqrt{2}$ for transverse oscillations[1]). However, using the value of the cathode current density for j Eqs. (11) and (2)give frequencies of the same order of magnitude as observed.

It has been shown[1] that plasma ion oscillations are excited by secondary electrons from electrodes surrounding the electron beam region. The excitation of the electron plasma oscillations is very likely due to electrons traveling in a circular orbit somewhere between the electron gun and the cone tips. Assume that electrons of a velocity corresponding to V volts travel in a transverse magnetic field B_0 in a circular orbit of radius r_0. Then

$$r_0=3.37\times10^{-6}(V^{\frac{1}{2}}/B_0) \tag{4}$$

and the related gyro frequency

$$f_g=2.8\times10^{10}B_0. \tag{5}$$

Combining (4) and (5) gives

$$r_0=9.43\times10^{4}(V^{\frac{1}{2}}/f_g). \tag{6}$$

Table I shows the results of measurements of the frequency f_e and beam voltage V. Also shown in Table I

TABLE I.

V volts measured	f_e Mc measured	f_g Mc calculated	r_0 mm calculated
310	332	332(measured)	5.00
330	342	342	5.01
343	352	353	4.96
360	360	364	4.97
380	368	374	4.99
410	380	385	5.02
440	397	395	4.98
470	407	406	5.02
510	418	417	5.08
500	420	427	5.02
550	437	437	5.06
570	455	448	4.95

is r_0 calculated from (6) putting $f_g=f_e$. Table I shows that r_0 is approximately constant equal to 5 mm. The electron gun is situated at a distance 14 mm from the center. The magnetic field strength could not be measured accurately but satisfied roughly (5) at $r=r_0$ for $f_e=f_e$. In Table I is also shown f_g calculated from (5) by putting $f_g=f_e$ for $f_e=332$ mc and assuming proportionality between B_0 and the measured current through the electromagnet.

From Table I it is seen that the measured frequency f_e is equal to both the electron plasma frequency and to the gyro frequency. It is, therefore, most likely that the observed oscillation is a phenomenon of resonance in a circular orbit at the radius r_0. Because of either longitudinal[3] or transverse electron plasma oscillations a standing electrostatic potential wave is presumably set up within the electron beam. The mechanism by which the electrons supply energy to the plasma field is not known.

[3] W. O. Schumann, Z. Physik **121**, 7 (1943).

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Power Flow in Electron Beams

L. R. WALKER
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey
(Received December 11, 1954)

The ac power flow in an electron beam moving in weak electric fields is given a precise definition. Since it is a quantity quadratic in the applied field strength its evaluation requires the quadratic perturbation theory of the beam to be worked out. Quadratic expressions in the first-order variables of the beam and circuit may also be written down which have the dimensions of power and are shown to satisfy a conservation theorem. The relations between the exact quantities of the nonlinear theory and the constructs from linear theory are discussed.

MOST of the analysis of problems concerned with the propagation of electromagnetic waves on electron beams uses the small signal or linearized form of the relevant equations. It is not always clear how to fit such typically quadratic quantities as the ac energy or ac power flow into the framework of this linear theory. There are cases in the literature where this has been done incorrectly.[1] These questions have recently been discussed for the case of an indefinitely extended beam of electrons interacting only with an external traveling wave electric field.[2] In such a system special difficulties arise which make it necessary to specify carefully how the ac field has been set up in order to obtain unequivocal results. In finite systems, where a beam enters a circuit in a known state of modulation, some of the earlier ambiguities are absent. However, even in this case, there appears to be a divergence of opinion over the most satisfactory procedure.[3] The difficulty appears to be, at least in part, a semantic one. In this paper an attempt is made to clarify ideas by considering in some detail the power flow in particular model representing an electron beam interacting with a circuit.

To exhibit the character of the problem suppose we have a beam of electrons whose charge density ρ and velocity (in the x direction) v are functions of x and t, the time, only. We consider only longitudinal motions and longitudinal electric fields. The density of kinetic energy flow is clearly $m\rho v^3/2e$ for such a beam at any point. If the applied fields are so weak that we are in the regime where the equations of the problem may be linearized and if these fields are periodic (in time), it is consistent with usual practice to define the density of rf power flow as that part of the time average of $m\rho v^3/2e$ which is of the second order in the applied field. We shall write this as $m[\rho v^3]_2^0/2e$ and consistently use superscripts to indicate the harmonic of the fundamental frequency and subscripts to denote the order (in the applied field).[4] The expression, $m[\rho v^3]_2^0/2e$, clearly contains, if expanded, such terms as $m[\rho]_2^0 v_0^3/2e$, which depend upon second-order quantities, in this case $[\rho]_2^0$. Such terms are as large as those depending upon products of first-order quantities, such as $3mv_0^2[\rho_1 v_1]^0/2e$, which may occur in the power flow. It is also clear that they will appear as the beam progresses even if the latter does not contain them when it initially enters the circuit.

The appearance of such explicitly second-order terms as $[\rho]_2^0$ might lead one to suppose that it would be necessary to carry out a solution of the quadratic equations. It may, therefore, be asked whether or not it is possible to re-express these second-order quantities in terms of first-order ones. If so, then one also can write the rf kinetic power flow in terms of expressions taken from linear theory. This proves to be the case. An alternative approach is to find quadratic expressions in the first-order quantities which have the dimensions of power flow or density of kinetic energy and which appear to satisfy those requirements, such as conservation laws, which may reasonably be asked of them. This is essentially the method used by Louisell and Pierce.[3] We shall attempt to exhibit here the relation between the two methods.

If we are given a specific circuit interacting with an electron stream we can in principle analyze the interaction by expanding in normal modes, or, if we see fit, go further to find a Green's function for a current element exciting the circuit. For our purposes we may use a model which incorporates those features of the general interaction which seem relevant. Thus, we shall assume that the fundamental component of the ac current excites a single degenerate pair of backward and forward modes; that the higher harmonics excite nothing; that the effect of all cut-off modes can be lumped together into a single "space-charge voltage, V_s which acts on the electrons. It is known[5] that for uniformly charged disk electrons on the axis of a conducting cylinder the space-charge force between "electrons" goes very nearly as $e^{-x/\lambda}$, where x is the

[1] On the other hand, the correct procedure was outlined very explicitly in an early paper by W. C. Hahn, Gen. Elec. Rev. **42**, 497 (1939). I am grateful to the referee for calling this paper to my attention.
[2] L. R. Walker, J. Appl. Phys. **25**, 615 (1954).
[3] See W. H. Louisell and J. R. Pierce, "Power flow in electron beam devices" (to be published in Proc. Inst. Radio Engrs.) for an approach different from that of this paper.
[4] Since this somewhat unfamiliar notation is used throughout, the reader's attention is particularly directed to it.
[5] Tien, Walker, and Wolontis, Proc. Inst. Radio Engrs. **43**, 260 (1955).

separation of the electrons and λ is a range parameter. This implies that for V_s we may write a generalized Poisson equation

$$-\lambda^2\frac{\partial^2 V_s}{\partial x^2}+V_s=\rho/C_1, \tag{1a}$$

where C_1 is a space charge parameter Clearly if we let $\lambda^2 C_1$ remain finite as λ becomes large we obtain the usual Poisson equation for "plane" electrons.

We have, in addition, the equations of continuity and of motion

$$\frac{\partial\rho}{\partial t}+\frac{\partial\rho v}{\partial x}=0, \tag{1b}$$

$$\frac{\partial v}{\partial t}+\frac{1}{2}\frac{\partial v^2}{\partial x}=-\frac{e}{m}\left[\frac{\partial V_s}{\partial x}+\frac{\partial V_c}{\partial x}\right], \tag{1c}$$

where V_c is the circuit voltage, the sum of the voltage amplitudes of the two circuit waves, V_c^+ and V_c^-.

We first consider the problem of writing down a plausible expression for power flow in the linearized theory. Suppose Eqs. (1a) to (1c) are linearized, $\partial/\partial t$ put equal to $j\omega$ and a factor $e^{j\omega t}$ suppressed; then we have, if

$$i_1=\rho_0 v_1+v_0\rho_1, \tag{2a}$$

the equations,

$$j\omega\rho_1+\frac{\partial i_1}{\partial x}=0, \tag{2b}$$

$$\left(j\omega+v_0\frac{\partial}{\partial x}\right)v_1=-\frac{e}{m}\left[\frac{\partial V_{s1}}{\partial x}+\frac{\partial V_{c1}}{\partial x}\right], \tag{2c}$$

$$-\lambda^2\frac{\partial^2 V_{s1}}{\partial x^2}+V_{s1}=\frac{\rho_1}{C_1}, \tag{2d}$$

where the subscripts "0" and "1" refer to zero and first-order quantities.

The excitations of the backward and forward circuit wave amplitudes, V_{c1}^+ and V_{c1}^-, are given by

$$\begin{aligned}\frac{\partial V_{c1}^+}{\partial x}+j\beta V_{c1}^+&=-\frac{Z_0}{2}\frac{\partial i_1}{\partial x},\\ \frac{\partial V_{c1}^-}{\partial x}-j\beta V_{c1}^-&=+\frac{Z_0}{2}\frac{\partial i_1}{\partial x},\end{aligned} \tag{2e}$$

where β is a cold propagation constant and Z_0, a circuit impedance. These equations are readily derived from the pair of usual transmission line equations by forming the appropriate linear combinations. It is now possible by a series of straightforward manipulations on Eqs. (2) to show that

$$\begin{aligned}\frac{\partial}{\partial x}\Bigg[&\frac{V_{c1}^+(V_{c1}^+)^*-V_{c1}^-(V_{c1}^-)^*}{Z_0}\\ &+\frac{(V_{c1}^++V_{c1}^-)^*i_1+(V_{c1}^++V_{c1}^-)i_1^*}{2}\\ &+\frac{V_{s1}^*i_1+V_{s1}i_1^*}{2}+\frac{mvo}{2e}(i_1^*v_1+i_1v_1^*)+\frac{j\omega\lambda^2C_1}{2}\\ &\times\left(V_{s1}\frac{\partial V_{s1}^*}{\partial x}-V_{s1}^*\frac{\partial V_{s1}}{\partial x}\right)\Bigg]=0,\end{aligned} \tag{3}$$

where the asterisk indicates a complex conjugate. The term in brackets has the dimensions of a power and is independent of distance; it should therefore contain all the identifiable ac power flows in the whole system. One may identify the terms as: (1) circuit power in the forward and backward waves, (2) rate of working against applied voltage, (3) rate of working against space charge voltage, (4) kinetic power flow, and (5) mutual potential energy of the space charge carried by the beam. By rearranging the terms of (3) in various ways one may produce other recognizable or nameable powers. But in whatever way this is done one might hope that the whole quadratic form in first-order quantities represented in (3) would be the total ac power flow.

We now consider the nonlinear equations. By a series of steps analogous to those used on the linearized equations, one finds the following relation:

$$\begin{aligned}\frac{\partial}{\partial t}\left[\frac{m\rho v^2}{2e}+\frac{C_1}{2}V_s^2+\frac{\lambda^2C_1}{2}\left(\frac{\partial V_s}{\partial x}\right)^2\right]+\frac{\partial}{\partial x}\Bigg[&\frac{m\rho v^3}{2e}\\ &+\rho vV_s-\lambda^2C_1V_s\frac{\partial^2V_s}{\partial t\partial x}\Bigg]=-\rho v\frac{\partial V_c}{\partial x}.\end{aligned} \tag{4}$$

The right-hand side is the rate of working of the fields due to the propagating circuit modes on the beam. It is reasonable then to identify the two bracketed terms with the total stored energy and total power flow in the beam, respectively. We are taking the viewpoint that the space charge terms are to be considered as a property of the beam even though their value may actually depend upon the nature of the circuit. The term $m\rho v^3/2e$ in the power is that which has already been identified with the kinetic power flow and the other two terms represent the total current (convection +displacement) associated with the beam multiplied by the space charge potential. The ac power flow in the beam, in the sense in which we agreed to use the term, is now

$$\left[\frac{m\rho v^3}{2e}+\rho vV_s-\lambda^2C_1V_s\frac{\partial^2V_s}{\partial t\partial x}\right]_2^0. \tag{5}$$

It is clear that the terms relating to the circuit are just the same as in the linear case since we have assumed that the circuit is only excited at the fundamental frequency.

Suppose we now average the equations of motion and continuity, (1b) and (1c), over time. Then we obtain

$$\frac{\partial}{\partial x}[\rho v]^0=0, \tag{6a}$$

and

$$\frac{\partial}{\partial x}\left[\frac{v^2}{2}+\frac{e}{m}V_s\right]^0=0. \tag{6b}$$

The circuit voltage V_c contributes nothing since it is supposed to be excited only by the fundamental. It may be noted from Eq. (1a), though we do not use the result, that

$$-\lambda^2\frac{\partial^2[V_s]^0}{\partial x^2}+[V_s]^0=\frac{[\rho]^0}{C_1}. \tag{7}$$

Now consider the various terms in the Eq. (5). We have

$$\left[\frac{m\rho v^3}{2e}\right]_2^0=\frac{m}{e}[\rho v]_2^0\left[\frac{v^2}{2}\right]_0^0 +\frac{m}{e}\left[[\rho v]_1^1\left[\frac{v^2}{2}\right]_1^1\right]^0+\frac{m}{e}[\rho v]_0^0\left[\frac{v^2}{2}\right]_2^0.$$

But from (6a) and (6b)

$$[\rho v]_2^0=0,$$

$$\left[\frac{v^2}{2}\right]_2^0=-\frac{e}{m}[V_s]_2^0+\text{constant independent of } x.$$

Thence,

$$\left[\frac{m\rho v^3}{2e}\right]_2^0=\frac{m}{e}v_0[i_1v_1]^0-i_0[V_s]_2^0+\text{constant}, \tag{8}$$

where, as before,

$$i_1=[\rho v]_1^1,$$

and

$$i_0=[\rho v]^0=\rho_0v_0.$$

Again

$$[\rho vV_s]_2^0=[\rho v]_0^0[V_s]_2^0+[[\rho v]_1^1[V_s]_1^1]^0+[\rho v]_2^0[V_s]_0^0 =i_0[V_s]_2^0+[i_1[V_s]_1^1]^0. \tag{9}$$

Finally

$$-\lambda^2C_1\left[V_s\frac{\partial^2V_s}{\partial t\partial x}\right]_2^0=-\lambda^2C_1\left[[V_s]_1^1\left[\frac{\partial^2V_s}{\partial t\partial x}\right]_1^1\right]^0.$$

Put

$$[V_s]_1^1=V_{s1}e^{j\omega t}+V_{s1}^*e^{-j\omega t}.$$

Then

$$\left[\frac{\partial^2V_s}{\partial t\partial x}\right]_1^1=j\omega\left(\frac{\partial V_{s1}}{\partial x}e^{j\omega t}-\frac{\partial V_{s1}^*}{\partial x}e^{-j\omega t}\right),$$

and

$$-\lambda^2C_1\left[[V_s]_1^1\left[\frac{\partial^2V_s}{\partial t\partial x}\right]_1^1\right]_2^0=-j\omega\lambda^2C_1 \times\left(V_{s1}^*\frac{\partial V_{s1}}{\partial x}-V_{s1}\frac{\partial V_{s1}^*}{\partial x}\right). \tag{10}$$

We may now form the expression (5) by adding the results (8), (9), and (10) and compare it with those terms of (3) which do not refer explicitly to the circuit. Remembering that if f and g are any two of the quantities in which we are interested, then $[[f]_1^1[g]_1^1]^0$ and $\frac{1}{2}[f_1^*g_1+f_1g_1^*]$ are simply two ways of writing the same quantity, it is clear that (5) and (3) agree exactly. This shows then that (3) does account for the whole of the ac power flow. The term by term comcomparison of (3) and (5) however indicates that $m[\rho v^3]_2^0/2e$ is not identical with $mv_0(i_1^*v_1+i_1v_1^*)/2e$, which was earlier identified with the kinetic power flow, as might perhaps have been hoped. Neither is $[\rho vV_s]_2^0$ equal to $\frac{1}{2}[V_{s1}^*i_1+V_{s1}i_1^*]$ which had been previously labelled as the work done by the convection current against the space charge force. The differences arise in each case from the existence of the term $i_0[V_s]_2^0$, which appears in the two expressions (8) and (9) with opposite and compensating signs. There appears to be no reason to doubt that this term has, in general, a value different from zero. As can be seen from Eq. (7), $[V_s]_2^0$ is excited by any change in the dc density of the beam from its unperturbed value and such changes certainly occur. The only case in which the status of this term might be in some doubt is that in which positive space charge due to ions was included in the calculations. Such a positive density would not affect any other part of the analysis but it seems reasonable to suppose that if there existed a dc potential gradient arising from the variation of $[V_s]_2^0$, then the ions would drift in such a way as to modify and, perhaps, annul $[V_s]_2^0$. Clearly, where one is interested in adding the power terms together, the $[V_s]_2^0$ term will drop out and in such cases it is permissible to proceed as though it did not exist.

On the Space-Charge-Limited Current between Nonsymmetrical Electrodes

GEORGE D. O'NEILL
Research Laboratories, Sylvania Electric Products Inc., Bayside, New York
(Received July 26, 1954; revised manuscript received May 5, 1955)

The validity of employing an equivalent symmetrical anode for expressing the space-charge-limited current between nonsymmetrical electrodes is examined. It is shown that when the equivalent anode is determined by the capacitance, the error in the expression for the current is usually quite small and its approximate magnitude may often be anticipated. Exact expressions for the capacitance of some basic electrode arrangements are given, followed by a brief resumé of corrections required for interpretation of experimental data.

I. INTRODUCTION

RIGOROUS solutions for the space-charge-limited current, with but one exception, have been derived only for electrode geometries in which the electron paths are straight lines, consequently, resort is had to approximation methods in many practical cases.[1–5]

Experience has shown that one of these method is, with but rare exception, so reliable that it may be stated as a general rule: *The space-charge-limited current between a cathode and an anode of arbitrary shape is approximately that which would be expected for the same cathode and a symmetrical anode when the space-charge-free capacitances are equal.*

Critical analysis of the general problem leads to the conclusion that the rule should hold to a highly useful degree of approximation provided an accurate knowledge of the cold capacitance can be obtained.

The purpose of this paper is partly to investigate the degree of accuracy that can be expected from computations, partly to show the kind of problem where serious departures from accuracy may be encountered and how they may be interpreted, and partly to provide accurate expressions for the capacitance of some electrode geometries into which many of the space-charge problems can be resolved. No attempt will be made to extend the results into specialized fields such as electron guns or hollow cathodes.

In Sections II and III the usual simplifying assumptions will be made. That is, the initial electron velocity is zero, the contact potential difference is zero, and the cathode is a unipotential conductor. In Section IV the simplest methods for bringing the necessary corrections for these effects into the computation will be outlined.

A list of symbols frequently used is given below:

A—area of an electrode.
i—current in amperes.
$V_0; V_a; V$—potential of cathode; of anode, at a point.
$E_0; E_x$—electric field intensity at surface of cathode when cold; at point x in x direction.
l—length of an electrode assembly.
$s; h$—distance between plane electrodes; from axis of cylindrical cathode to plane anode.
r_e—radius of coaxial cylinder for equal capacitance.
a—distance between axes of cylinders.
$b; c$—distance of line charge from axis; from axis to mirror image of induced line charge.
q—charge per unit length in cgs units.
$Q; Q'$—electrostatic charge V_aC; space charge iT when $E_0=0$, $V_a\neq 0$.
$C; C'$—capacitance in vacuum dielectric; quotient of space charge by anode potential, sometimes referred to as "hot capacitance."
$r_0; r; r_a$—radius of cathode; a varying radius; radius of cylindrical anode.
$k_1; k_2; \beta$—2.334 microamperes/volt$^{\frac{3}{2}}$; $2\pi k_1$; a dimenless quantity used in Langmuir's equation.
$e; m$—charge of an electron; rest mass of an electron.
$T; T_m$—transit time of an electron when current is space-charge limited; mean transit time for all electrons in the system.

II. CONCERNING THE VALIDITY OF THE RULE

The idea that the space-charge relations for electrode configurations other than those for which rigorous solutions have been derived can be usefully approximated by relating the capacitance to that of a symmetrical pair contains a certain amount of logic. One can take the expressions for the current between symmetrical electrodes as a starting point and determine in a general way the result of introducing variations.

For parallel planes, Child's equation for the current is

$$i=(AV_a^{\frac{3}{2}}/9\pi s^2)(2e/m)^{\frac{1}{2}}, \tag{1}$$

which in practical units is

$$i=k_1V_a^{\frac{3}{2}}/s^2 \tag{1a}$$

where k_1 has the value 2.334×10^{-6} ampere/volt$^{\frac{3}{2}}$.

[1] G. D. O'Neill, Sylvania Technologist **III**, 22 (1950).
[2] G. D. O'Neill in Henney's *Radio Engineering Handbook* (McGraw-Hill Book Company, Inc., New York, 1950), fourth edition, Chap. 13.
[3] W. G. Dow, *Fundamentals of Engineering Electronics* (John Wiley and Sons, Inc., New York, 1952), second edition, pp. 158–162. (Improved presentation of material in reference 1.)
[4] H. F. Ivey, J. Appl. Phys. **24**, 1466–1472 (1953).
[5] M. Matricon and S. Trouvé, L'Onde Elec. **30**, 510–521 (1950). (The approach is different from that employed in reference 1 and this paper.)

The capacitance is

$$C=A/4\pi s \tag{2}$$

in electrostatic units; in practical units it is

$$C=1.112\times10^{-12}A/4\pi s \tag{2a}$$

farads when the dimensions are in centimeters.

With the electrodes cold, the magnitude of the charge on either electrode is $Q=V_aC$ and the field $E=-V_a/s$. When the cathode is passing a space-charge-limited current, the field at the cathode is zero and at the anode it is increased by the factor 4/3, as may be seen by a single integration of Poisson's equation or by taking dV/ds from Eq. (1) with i constant. Therefore we may take as the "hot capacitance" four-thirds the value given by Eq. (2): $C'=A/3\pi s$. Also, since the field at the cathode is zero, its charge is zero; consequently, the total charge on the anode is equal in magnitude to the total space charge in the interelectrode region. One thus has $Q'=C'V_a=iT$, where T is the electron transit time,

$$\begin{aligned}T&=V_a(A/3\pi s)[9\pi s^2/AV_a^{\frac{1}{2}}(2e/m)^{\frac{1}{2}}]\\&=3s/(2V_ae/m)^{\frac{1}{2}},\end{aligned} \tag{3}$$

in which the denominator has the value $5.93\times10^7V_a^{\frac{1}{2}}$ cm/sec.

It can be shown easily that if V_a and s are varied so that any pair of values V_a' and s' yield the original field $E_0=-V_a/s$ at the cathode, the space charge Q' also remains constant. Now if the anode is in some way altered so that the symmetry is destroyed while C is unaltered, one may take as a working hypothesis the idea that the total space charge required to reduce E_0 to zero over the entire cathode surface remains substantially unchanged. To this extent, the product iT_m, in which T_m is the mean transit time, also is substantially unchanged. The meaning of *substantially* will later be made evident. Obviously T_m is apt to be less than T, since the transit time for an electron through an element of cathode surface varies approximately with the distance, while the current density varies approximately with the reciprocal of distance squared. The rule does not include a correction for the difference between T for symmetrical electrodes and T_m for nonsymmetrical electrodes, consequently, the predicted current will be in error by the factor T_m/T, which in most cases means that it will be too low.

When the cathode is a circular cylinder the error will be less than when it is a plane for reasons that will will become evident. For coaxial cylinders, Langmuir's equation for the current is

$$i=(2lV_a^{\frac{3}{2}}/9r_a\beta^2)(2e/m)^{\frac{1}{2}} \tag{4}$$

or, in practical units,

$$i=k_2lV_a^{\frac{3}{2}}/r_a\beta^2 \tag{4a}$$

in which $k_2=14.65\times10^{-6}$ ampere/volt$^{\frac{3}{2}}$, and

$$\beta=\gamma-0.4\gamma^2+0.0917\gamma^3-0.0142\gamma^4\cdots, \tag{5}$$

where $\gamma=\ln(r_a/r_0)$. The cold capacitance is

$$C=l/2\ln(r_a/r_0). \tag{6}$$

The field at the cold cathode is $2Q/r_0$ where $Q=CV_a$, thus $E_0=-V_a/r_0\ln(r_a/r_0)$. With the cathode heated so that E_0 goes to zero, the charge induced on the anode is the space charge $Q'=iT$. Taking i constant in Eq. (4) so that the field at the anode can be determined, one has

$$Q'=(V_a/3)l[1+(2r_a/\beta)d\beta/dr_a]. \tag{7}$$

Correspondingly, the transit time T is the quotient of Eq. (7) by Eq. (4):

$$T=\frac{(3/2)r_a[\beta^2+2r_a\beta(d\beta/dr_a)]}{(2V_ae/m)^{\frac{1}{2}}}. \tag{8}$$

As was the case for parallel planes, any value of E_0 requires the same value of space charge Q' per unit length of the system if E_0 is to be reduced to zero when the cathode is heated, *regardless of anode radius* r_a. That is, if r_a is changed and V_a adjusted so that E_0 is brought back to its original value, Q' remains unchanged.

Consider now two pairs of coaxial cylinders, both having the same cathode radius r_0. Let the lengths be l_1 and l_2, the anode radii r_1 and r_2, and the capacitances C_1 and C_2. At a potential V_a the space charge will be Q_1' and Q_2'. Now for a pair of coaxial cylinders in which the cathode has the same radius r_0, but whose length is l_1+l_2 and capacitance C_1+C_2, the radius r_e can be determined by substituting $r_e\equiv r_a$ in Eq. (6). When this radius, along with the values of β and $d\beta/dr_e$ for $\gamma=\ln(r_e/r_0)$ are substituted in Eq. (7), it is found that $Q'=Q_1'+Q_2'$ for the same value of V_a.

These observations lead to the same hypothesis that was drawn from the discussion of planes, i.e., if the symmetry of the system is destroyed by altering the anode, Q' or C' must remain substantially the same if C is not altered.

Referring again to Eq. (7), it may be noted that as r_a/r_0 increases, β as given by Eq. (5) approaches unity and $d\beta/d\gamma$ or $d\beta/dr_a$ approaches zero. It follows that, for constant V_a, Q' approaches a constant value as r_a/r_0 increases. Equation (4) shows that the current-anode radius relation becomes linear under this condition, while Eq. (8) shows that the transit time-anode radius relation also becomes linear. From this it follows that where the dissymmetry with cylindrical cathodes is no more extreme than would normally be encountered in practical cases, the mean transit time T_m for the nonsymmetrical pair will not differ appreciably from the transit time T for the symmetrical pair when the capacitances and anode voltages are equal, and that the error involved in applying the rule will be very small.

Beyond the generalities mentioned, a number of problems can be designed for testing the rule. An example of these will be discussed, and it will become

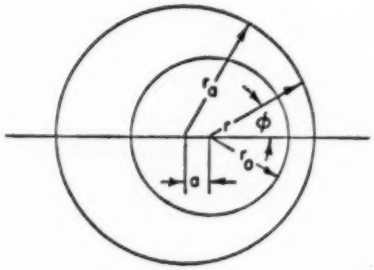


FIG. 1.

evident that the results can be extended to common electrode configurations.

In this example both cathode and anode are circular cylinders, their radii being r_0 and r_a, respectively. Their axes are parallel but displaced a distance a as shown in Fig. 1. By choosing $r_0=100$, $r_a=102$, and $a=1.0$ cm, the spacing on the right in the figure is one-third the spacing on the left, hence the current density will be about nine times greater and the transit time about one-third as great.[6]

In the figure, the spacing between cathode and anode is $r-r_0$, consequently, we have from Eq. (1) $\Delta i\simeq k_1 l V_a^{\frac{1}{2}} \times\Delta\varphi/(r-r_0)^2$, and from Eq. (4) $\Delta i\simeq k_2 l V_a^{\frac{1}{2}}\Delta\varphi/r\beta^2$. The former expression is somewhat less accurate but easier to evaluate and will be used in obtaining the desired approximation; the difference between this and the one containing β is only about 0.5% for $a=0$.

By inspection of the figure, $r_a^2=(r\sin\varphi)^2+(a+r\cos\varphi)^2$. Taking $r_a^2\gg a^2\sin^2\varphi$, the problem is simplified by dropping the small term, and one has

$$i\simeq 2k_1 V_a^{\frac{1}{2}} r_0 l\int_0^{\pi}\frac{d\varphi}{(r_a-r_0-a\cos\varphi)^2},$$
$$\simeq\frac{k_2 V_a^{\frac{1}{2}}(r_a-r_0)r_0 l}{[(r_a-r_0)^2-a^2]^{\frac{1}{2}}}. \tag{9}$$

For the dimensions given, Eq. (9) yields $i\simeq 563$ microamperes; a numerical solution based on Eq. (4) yielded a figure about one percent higher. A simpler correction is to use the ratio of the two solutions shown for $a=0$, in which i by the parallel plane expression was 366 and by the coaxial expression it was 368: the result is $i\simeq 563\times 368/366=567$. When one considers the fact that the lines of electric force with these dimensions are nearly straight, it seems likely that a solution rigorously obtained would probably be not more than a small fraction of one percent greater.

The capacitance for this electrode arrangement is[7]

$$C=\frac{l}{2}\left[\cosh^{-1}\left(\frac{r_a^2+r_0^2-a^2}{2r_a r_0}\right)\right]^{-1}. \tag{10}$$

[6] The dimensions given for r_0, r_a, and a are obviously of a different order of magnitude from those of practical electrodes and were chosen merely for convenience. If these and length l are multiplied by 10^{-3}, the cathode radius is then one millimeter, but the current and capacitance per unit length remain unchanged.

[7] William R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company, Inc., New York, 1950), second edition, pp. 76–78.

Tables of hyperbolic functions for the solution of Eq. (10) to the necessary degree of accuracy for the dimensions used in the present example are not always available; for this reason, and because reference to the field at the cathode will be made in the discussion to follow, a different expression for the capacitance will be shown.

In Fig. 2, a line charge q per unit length is placed inside the cathode at b, and its mirror image in both cylinders at c. One has

$$bc=r_a^2 \tag{11}$$
$$(b-a)(c-a)=r_0^2 \tag{12}$$

for which

$$c=(1/2a)[a^2+r_a^2-r_0^2+[(a^2+r_a^2-r_0^2)^2-4a^2r_a^2]^{\frac{1}{2}}] \tag{13}$$

and b can be had by putting the result from Eq. (13) back into Eq. (11). The field at a point on the x axis is then

$$E_x=-2q[1/(x-b)+1/(c-x)] \tag{14}$$

in which q is the charge per unit length, CV_a/l, when the potential of the cathode is zero. Integration of Eq. (14) between r_0+a and r_a yields the potential V_a at r_a; one then has $C=V_a/q$, which is

$$C=l\left\{2\ln\left[\frac{(r_a-b)(c-r_0-a)}{(c-r_a)(r_0+a-b)}\right]\right\}^{-1}. \tag{15}$$

The equivalent radius r_e is then obtained by replacing r_a by r_e in Eq. (4) and equating to Eq. (15).

Taking $r_0=100$, $r_a=102$, $a=1.0$ as before, one has $r_e=101.73$, and when this value is used in place of r_a in Eq. (4) and Eq. (5), the current for $V_a=1.0$, $l=1.0$, is 500 microamperes. Assuming the true value of the current to be about 567 microamperes, use of the rule in this problem is seen to result in an error of 11.8%.

This considerable discrepancy may at first seem quite disturbing. As a matter of fact, the dimensions were purposely chosen so that the dissymmetry would be so extreme that a reliable figure for the error could be determined. It will be obvious that (a) substantially all of the error is due to the fact that the rule makes no allowance for the difference between the mean transit time in the nonsymmetrical arrangement and the transit time for its symmetrical equivalent, and (b) the error

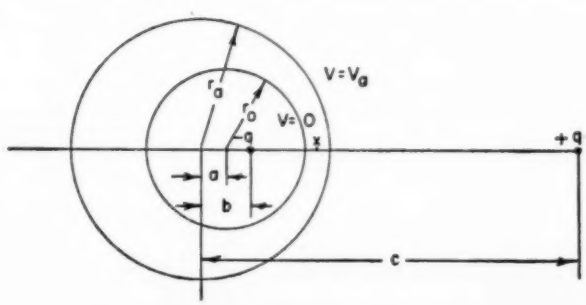


FIG. 2.

decreases to negligibly small values for electrode geometries one encounters in practice.

The capacitance of the symmetrical and nonsymmetrical electrode arrangements is 29.1 esu by Eq. (15) for $l=1$ cm. Since r_a/r_0 and r_e/r_0 are so near unity, the space charge can be taken with negligible error as $Q'=(4/3)\times 29.1\times 1.112\times 10^{-12}=4.32\times 10^{-11}$ coulomb for both at unit voltage. Assuming the true current to be 5.67×10^{-4} ampere, the mean transit time T_m for the nonsymmetrical diode is 7.62×10^{-8} second while for the symmetrical diode, in which the current was found to be 5.0×10^{-4} ampere, the transit time T is 8.64×10^{-8} second.

If now the dissymmetry is slightly decreased, it can be seen that the error decreases quite rapidly. For example, let r_a be changed from 102 to 104 cm while all other dimensions are as before. The ratio of maximum-to-minimum spacing and transit time has dropped to (5/3) and the ratio of maximum-to-minimum current density has dropped to about $(5/3)^2$, consequently, the error is much lower. By Eq. (9), $i\simeq 104$ microamperes; by Eq. (15), $C=13.21$ statfarads, and this value in Eq. (6) yields 103.85 for the equivalent anode radius. Solving for the current by Eq. (4), $i\simeq 102.5$ microamperes, an error of less than 2%.

Considering that the field intensity E_0 at the surface of the cathode when cold, at a given value of V_a, is a fair measure of the current density when the cathode is hot, one can relate in a general way the spread in E_0 to the difference between T and T_m. Taking dimensions that might resemble a practical situation, let $r_0=1.0$, $a=50$, $r_a=100$, $l=1.0$. One can use Eq. (14) to find the minimum and maximum values of E_0 by solving for E_x at $x=b+r_0$ and $b-r_0$, respectively. For the present example, it is sufficiently accurate to take $b=a$ and $c=r_a^2/b$, and one has for the maximum $E_0=-2q(1+1/149)$ and for the minimum $E_0=-2q(1+1/151)$. The ratio of these two values is so near unity one concludes that the current density around the periphery of the cathode is almost constant. In view of the fact that the space charge Q' at constant electrode voltage approaches a constant value as r_a/r_0 becomes large, it appears that the mean transit time in the present example is very nearly the same as in the symmetrical equivalent.

Analysis of some other hypothetical cases in which cylindrical cathodes were used substantiate the conclusion that can be drawn from the example discussed here, namely, that the simple method described for computing the space-charge-limited current is about as accurate as the methods normally available for measuring the result. When the cathode is taken as a plane and the anode is a warped surface or a plane not parallel to the cathode, the rule appears to apply with useful accuracy in cases where the ratio of maximum-to-minimum spacing between the electrodes is not greater than about three.

III. CAPACITANCE RELATIONS FOR BASIC ELECTRODE CONFIGURATIONS

In practice, most problems involving the space-charge-limited current where the cathode is a cylinder and the electrons follow curved paths can be reduced to one of several basic forms. In view of the useful accuracy with which the rule appears to apply in most cases, it is worthwhile to employ expressions for the capacitance that are as accurate as can be had. Most of the simpler expressions found in handbooks disregard small quantities, consequently, errors in the computation of current may be considerable if they are employed. In this section most of the more important capacitance relations for electrode structures in which the cathode is a cylinder will be given. Several of them appear in the list of approximation formulas given by Ivey,[4] but in some cases the results differ near the bounds.

It should be noted that certain electrode arrangements are entirely unsuited to the rule, especially those in which some of the electrons return to the cathode. For example, suppose the cathode and anode are both cylinders external to each other. Because of momentum, electrons tend to cut across the lines of force, and some may miss the anode completely and return to the cathode; in returning they have just as much influence on the space charge as those passing toward the collector. The case of a triode with positively-biased grid and negatively-biased plate is familiar; the current collected will be less, with the same total space charge, than when both grid and plate are positive simply because the current in the space-charge equations must be taken as the sum of the current in both directions.

1. Eccentric Cylinders

Two expressions for the capacitance have been given: Eqs. (10) and (15), the former being in general easier to use except for r_a/r_0 very close to unity.

2. Cylinder Parallel to a Plane

The derivation is given by Smythe,[7] the result being

$$C=l[2\cosh^{-1}(h/r_0)]^{-1}, \tag{16}$$

in which h is the perpendicular distance from the axis of the cylinder to the plane. A rather tedious analysis showed the error in using the rule to be negligibly small for $h/r_0>3.0$.

3. Cylinder between Infinite and Parallel Planes

In most tubes using a cylindrical cathode and an anode of rectangular cross section, the spacing between the cathode and the long sides of the rectangle is ordinarily so much lower than to the short sides that no appreciable error results if the cathode is taken to be located midway between a pair of infinite parallel planes whose separation is $2h$.

Expressions for the capacitance of this arrangement are given in several places. In their derivation small

terms are discarded which do not influence the result for large values of h/r_0, but as h/r_0 decreases toward unity, these simplified expressions break down seriously. In view of the commercial and experimental importance of this basic electrode arrangement in which h/r_0 may be quite small, it is worthwhile to obtain some improvement in the accuracy with which the capacitance or equivalent radius may be determined.

By dividing the cathode charge into two line charges placed in the plane through the axis and normal to the anode planes, and setting up an infinite array of images, a series expression for the capacitance can be derived that is accurate for all values of h/r_0. For low values of h/r_0, the series converges very slowly and computation is laborious, but it suggests a simple empirical formula that permits the computation of r_e/h to an accuracy of about two percent at $h/r_0=1.1$, the accuracy increasing as h/r_0 increases. It is:

$$C=l[2\ln[(4h^2-r_0^2/2)\pi hr_0]]^{-1}. \tag{17}$$

For $h\gg r_0$, Eq. (17) reduces to the expression usually given:

$$C=l[2\ln 4h/\pi r_0]^{-1} \tag{18}$$

and one has for this case $r_e/h=1.272h/r_0$.

It was stated that reasonably accurate prediction of the current can be expected for the electrode arrangement employing a cylindrical cathode parallel to a single plane anode when $h/r_0>3.0$. With the cathode midway between two parallel planes, acceptable accuracy should be had with even lower values of this ratio for the reason that the variation in E_0 around the circumference is less when V_a and h/r_0 are the same. Tests made by the writer some years ago on this arrangement yielded measured values of current that checked predicted values within a few percent for $h/r_0\sim 2.2$ when the effects of initial electron velocity and contact potential difference were taken into account

The effect of moving the cathode off center is occasionally of interest, especially in the investigation of microphonism. In such instances the cathode is usually a wire, so that $r_0\ll h$, and one has for the capacitance[8]

$$C=l[(4h/\pi r_0)\sin\pi(h-a)/2h)]^{-1} \tag{19}$$

in which a is the displacement of the filament in the direction of an anode plane.

4. Two Cylindrical Cathodes in Cylindrical Anode

Two cathodes of radius r_0 are placed diametrically opposite each other a distance a from the axis of the anode, the radius of the anode being r_a. The problem is set up in the same way as for the single off-center cathode in a cylindrical anode, except that line charges now appear at $-b$ and $-c$ as well as at b and c; their locations are described by Eqs. (11), (12), and (13). The field at the x axis due to all four charges is integrated between $x=a+r_0$ and $x=r_a$ to yield V_a when the potential of the cathodes is taken as zero. Each line charge is of magnitude $q=CV_a/2$, consequently, the capacitance is

$$C=l\left[\ln\left(\frac{r_a^2-b^2}{c^2-r_0^2}\right)\left(\frac{c^2-(r_0+a)^2}{(r_0+a)^2-b^2}\right)\right]^{-1}. \tag{20}$$

Where $r_a\gg r_0$, the line charges in the cathodes approach the cathode axes, so $b\to a$ and $c\to r_a^2/a$; Eq. (20) then comes down to

$$C=l\left[\ln\left(\frac{r_a^2-a^2}{(r_a^4/a^2)-r_0^2}\right)\left(\frac{(r_a^4/a^2)-(r_0+a)^2}{r_0^2+2r_0a}\right)\right]^{-1}. \tag{21}$$

In applying the rule to a result such as Eq. (20) or Eq. (21), it should be noted that since there are two cathodes the charge per unit length at any voltage under the electrostatic condition is only half as much as for the structure having the same capacitance but with only one cathode. Therefore if the average field intensity at the cathode surfaces is to be the same in both cases, it is necessary to take the length of the equivalent single-cathode structure to be twice as long as for the original, but with the same total capacitance. It follows that in applying the rule one must equate the capacitance of the actual structure to $C=2l/2\ln\times(r_e/r_0)$ and solve for the current in a coaxial diode whose length is $2l$, anode radius r_e and cathode radius r_0.

5. Cylindrical Anode with n Cylindrical Cathodes

A cylindrical anode of radius r_a has n cylindrical cathodes of radius r_0 uniformly disposed in a circle of radius a; the length of the anode and of each cathode is l.

There are a number of approaches to the solution for the capacitance, one of the simplest being to take the grid-plate capacitance of a cylindrical triode, then let the cathode radius go to zero and take the grid-plate capacitance under this condition.

Abraham[9] gave the grid plate capacitance as

$$C_{gp}=\frac{l\ln(r_a/r_c)}{2p[\ln(r_a/r_c)+\mu\ln(a/r_c)]}, \tag{22}$$

in which r_c is the radius of the cathode at the axis of the system and μ is the amplification factor. For the present problem we can take

$$p=(1/n)\ln(a/nr_0),$$
$$\mu=(p/n)\ln[(a/nr_0)(1-a^2/r_a^2)].$$

Taking the limit of Eq. (22) as $r_c\to 0$, one has $C=1/2p$

[8] P. H. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II, p. 1241.

[9] Max Abraham, Arch. Electrotech. 8, 42 (1949).

$\times(\mu+1)$, which is

$$C=\frac{nl}{2n\ln(r_a/a)+2\ln[(a/nr_0)(1-a^2/r_a^2)]}. \quad (23)$$

In applying the rule, Eq. (23) is equated to $C=nl/2\ln \times(r_e/r_0)$ and the current is computed for coaxial cylinders whose radii are r_0, r_a, and length nl.

Matricon and Trouvé,[5] on the basis of the simplifying assumption that "the form of the equipotential surfaces is the same under electrostatic and space charge conditions, and only the value of potential attached to each of these surfaces differs between the two conditions," developed an expression for the current applying to the present case. This expression was shown by Ivey[4] in the form yielding r_e/r_0, and it can be seen that the results differ considerably from those obtained by Eq. (23) at low values of n.

6. Cylindrical Anode with Helical Cathode

The cathode of radius r_0 is formed into a helix of radius a. If the separation between turns is large compared with r_0, the error in computing the capacitance is small if the unwound length L of the helix is made equal to the total length of n cathodes of radius r_0 equally spaced in a circle of radius a. Taking $L=nl$, in which l is the length of the electrode structure, substituting in Eq. (23) yields

$$C=\frac{L}{(2L/l)\ln(r_a/a)+2\ln[(al/Lr_0)(1-a^2/r_a^2)]}. \quad (24)$$

In applying the rule, Eq. (24) is equated to $L/2\ln(r_e/r_0)$ and the current is computed for a coaxial cylinder of length L.

7. Folded Filaments

In tubes using wire or small ribbon filaments folded as an M or a V and placed between rectangular grids or anodes, the capacitance is computed as though a single wire were used in a system whose length is the emitting length of the filament. Even with ribbon filaments, the spacing is usually large compared with the dimensions of the cathode so that the simple relation $r_e=1.27h$ is satisfactory for use in predicting the current.

IV. PRACTICAL CONSIDERATIONS

In the preceding sections the simplifying assumptions usually made in discussions of the three-halves-power law have been employed. In most applications one cannot afford to disregard such effects as those produced by initial electron velocity and the nonuniformity in potential of an equivalent anode in an amplifier tube. From a practical point of view, the development of a complete theory would probably be a rather fruitless undertaking for the reason that one rarely knows the values of such factors as the temperature and emission capability of all parts of a cathode. Nevertheless, there are available approximation methods which, with the aid of preliminary experimental data, can be employed with highly satisfactory results. A few of these, which the author has found to be particularly useful in bridging the gap between the simplified theory and its application, will be outlined.

1. Nonsymmetrical Grids

In amplifier tubes employing filamentary cathodes, the current density at the cathode surface is, except where the cathode-grid spacing is small compared with the spacing between grid wires, relatively uniform around the filament. Except at very low anode voltage, correction for such factors as voltage drop along the filament are adequately described in the existing literature and need not be reviewed here. One may, however, be warranted in pointing out what has proved to be a highly satisfactory procedure in handling design problems posed when the cathode diameter is large compared with the spacing between its surface and the grid. In such tubes the grid-cathode spacing, in a plane normal to their axes, is likely to vary considerably when measured at various points around the circumference. When the spacing between grid and anode in a triode or first and second grids in a multielectrode tube also varies circumferentially, the amplification factor, equivalent diode voltage, and current density will vary accordingly.

Expressions for the amplification are ordinarily derived for the electrostatic condition wherein μ is defined as C_{gk}/C_{pk}, the subscripts g, k, and p referring to grid, cathode, and plate, respectively. This is equivalent to defining μ by taking $dE_0/dV_g-\mu dE_0/dV_p=0$. Electronically, one has for the definition $di/dV_g-\mu di/dV_p=0$, and the two μ's are the same *only under certain ideal conditions*. If μ varies circumferentially or longitudinally, E_0 varies also. It follows that the electronic μ will be lower than the electrostatic μ due to the fact that the current density through regions of low μ (or high E_0) is ordinarily greater than through regions of high μ.

In the investigation of such a problem it is convenient to divide an enlarged cross section of the tube into a number of sectors so that the amplification factor, current, and transconductance of each sector can be computed separately on the basis of a simple, symmetrical system. The sums of the currents and transconductances of all the sectors then characterize the tube.[10] It may be noted in passing that the amplification factor is not the average value for the sectors; in a triode it is approximately $2g_mV_p[3i-2g_mV_g]^{-1}$, in which V_g and V_p are the grid and plate voltages, and g_m and i are the transconductance and plate current for the entire tube.

[10] G. Liebman, J. Inst. Elec. Engrs. (London) **93**, 138–152 (1946); see also G. R. Partridge, Proc. Inst. Radio Engrs. **37**, 87–94 (1949), for application to power tubes.

For a tube employing a variable-pitch grid, a satisfactory method of analysis consists in dividing the structure into slices perpendicular to the cathode as well as into sectors.

Although division of an amplifier tube into sectors for purposes of analysis or design is quite simple, there are cases in which it need not be used. For example, Deb and Sanyal[11] computed the characteristics of a triode in which the cross sections of the three electrodes constitute a family of confocal ellipses and obtained a reasonable result.

2. Initial Velocity and Contact Potential Difference

Of special importance, particularly in electron tubes operating at low electrode voltages, are the effects of initial electron velocity and the average contact potential difference between the cathode and the other electrode or electrodes. Initial velocity influences the apparent geometry of the device and adds to the electrode voltage;[12,13] the contact potential difference adds directly to the electrode voltage with no direct influence on the apparent geometry. It is not easy, in most cases, to separate completely the two influencing factors, consequently, they are combined in such a way that useful information can still be obtained, and one writes for the current,

$$i=G(V_a+\epsilon')^{\frac{3}{2}}, \tag{25}$$

in which G is the perveance and ϵ' is the correction voltage. In small electron tubes with coated cathodes, ϵ' is usually between 0.1 and 1.0 volt.

The perveance corresponds, in general, to the geometrical factors previously discussed, but includes the effect of initial electron velocity in moving the potential minimum away from the cathode. For example, one might take the perveance of a parallel plane diode to be $G=k_1A(s-s_{\min})^{-2}$, in which $s_{\min}$ is the distance between the cathode and the potential minimum.

The correction voltage ϵ' includes primarily the difference in potential between the cathode and the potential minimum, and the difference between the work function of the cathode and other electrode or electrodes. Substantially,

$$\epsilon'=\varphi_k-\varphi_a-V_{\min} \tag{26}$$

in which φ_k and φ_a are the average or effective values of work function for the cathode and anode (or a composite anode) over their surfaces, and $V_{\min}$ is the potential at the potential minimum with respect to the cathode. Both G and ϵ' vary with cathode temperature and with V_a, but this variation is usually small enough so that useful values can be determined from the static current and dynamic resistance measured at the same value of V_a. One has $di/dV_a=1/r_p$, so from Eq. (25) are obtained

$$G\simeq(2/3)^{\frac{3}{2}}/(r_p{}^3i)^{\frac{1}{2}}, \tag{27}$$

$$\epsilon'\simeq(3/2)ir_p-V_a. \tag{28}$$

For an amplifier tube, V_a in Eqs. (25) and (28) is replaced by $V_a=(\mu V_g+V_p)/(\mu+1)$, in which V_g is the potential (negative) of the first grid and V_p the potential of all other electrodes except cathode connected together, i is the total current, and μ the triode amplification factor. The total conductance is then $g_m(\mu+1)/\mu$; substituting in Eqs. (27) and (28), one has the corresponding values for the amplifier tube:

$$G\simeq(1/i)^{\frac{1}{2}}[(2/3)g_m(\mu+1)/\mu]^{\frac{3}{2}}, \tag{27a}$$

$$\epsilon'\simeq\frac{(3/2)i\mu}{g_m(\mu+1)}-\frac{\mu V_g+V_p}{\mu+1}. \tag{28a}$$

SUMMARY

The space charge $Q'=iT$ at a voltage V_a of a nonsymmetrical pair of electrodes is nearly the same as for a symmetrical pair if their electrostatic charges $Q=CV_a$ are equal and the area of the emitting electrodes of both are equal. From this it follows that if the mean transit time T_m in the nonsymmetrical arrangement is nearly the same as the transit time T for the symmetrical pair, then the space-charge-limited currents i are nearly equal.

Analysis shows that the method for predicting the current based on this assumption is reliable to a useful degree of accuracy. In some cases the error involved in unrestricted application of the rule can be determined in advance and suitable adjustment made.

Capacitance formulas for electrode systems basic to most of the electrode geometries have been given, and a separate section describes in outline the principal correction factors required for proper interpretation of the experimental data.

ACKNOWLEDGMENTS

The writer is indebted to W. G. Dow, E. R. Harrison, S. W. Harrison, L. A. Harris, L. S. Nergaard, and P. A. Redhead for constructive criticism of the draft, and to T. M. Liimatainen for laboratory assistance prior to World War II in verifying the rule experimentally.

[11] S. Deb and G. S. Sanyal, J. Appl. Phys. **25**, 1196–1203 (1954).
[12] I. Langmuir, Phys. Rev. **21**, 419–435 (1923).
[13] W. R. Ferris, RCA Rev. **10**, 134–149 (1949).

Correlation of Diffraction and Transmission Experiments for X-Ray and Neutron Elastic Scattering*

D. T. KEATING, *Brookhaven National Laboratory, Upton, New York*

AND

J. J. ANTAL,† *OMRO, Watertown Arsenal, Watertown, Massachusetts*

(Received February 23, 1955)

The correlation of a neutron cross-section plot in the cut-off region and a diffraction pattern has been considered. The cross section for elastic scattering by a sample as determined from transmission experiments has been related to its diffraction pattern. A cross-section plot can be transformed into a diffraction pattern by combining the cross section and its derivative. Examples of neutron cutoffs exhibiting strain and particle size broadening (carbonyl iron) and exhibiting particle size broadening alone (MgO) have been transformed. The patterns derived by the relation show good agreement with the x-ray diffraction patterns. Cross-section plots are sensitive to the state of crystalline perfection of the sample to the same degree as are diffraction patterns. Comparisons between the two should be made by performing the transformations given in the paper.

I. INTRODUCTION

DIFFRACTION effects are usually studied by direct observation of the scattered intensity. However, it is possible and with neutrons feasible, to study diffraction effects by the diminution of the incident beam in passing through the sample. After suitable corrections for absorption and inelastic scattering the transmission is determined by the diffraction of the incident beam. In a neutron transmission experiment, monochromatic neutrons are passed through a powder sample. The transmission, T, is measured and the cross section, σ, calculated according to

$$N\sigma = -\log T. \tag{1}$$

N is the number of atoms or molecules per cm^2. The capture and inelastic cross sections are approximately proportional to the neutron wavelength, Λ.[1] That part of the cross section arising from diffraction effects is also wavelength dependent showing a sawtooth-like pattern. As Λ is decreased from a value exceeding twice the largest spacing, scattering from this spacing suddenly appears with a corresponding increase in cross section. The scattered intensity and cross section from this set of planes decreases as Λ^2. Another jump appears when scattering occurs from the next largest spacing. The cutoffs occur when $\Lambda = 2d$, d being a spacing in the crystal.

The general features of such a plot have been explained by Halpern, Hamermesh, and Johnson[2] and others.[3] However, in explaining these features the diffracting material was assumed to be perfectly crystalline. Actually the cutoffs are not discontinuous, but have a detailed profile which is sensitive to the state of crystalline perfection of the sample. Neutron transmission techniques have recently been developed to such an extent that these diffraction effects can be studied in detail. Weiss[4] has made a study of neutron cutoffs with sufficient resolution to discuss the effects of strain upon the cutoffs of cold worked brass. The following is a generalization to include the effect of crystalline imperfection upon the profiles of the cutoffs.

II. THEORY

Figure 1 pictures a beam of neutrons with wavelength Λ incident on a powder sample. Some of the neutrons are scattered through the angle 2θ in the direction of the unit vector $\bar{s}$. The transmitted beam continues in the direction of the unit vector $\bar{s}_0$. The elastic scattering into the direction 2θ is determined by the interference function for a powder sample,[5]

$$N\bar{I}\left(\frac{S}{\Lambda}\right) = \sum_j f_j^2 + \sum_{j \neq k} f_j f_k \frac{\sin 2\pi (S/\Lambda)(\gamma_{jk})}{2\pi (S/\Lambda)(\gamma_{jk})}. \tag{2}$$

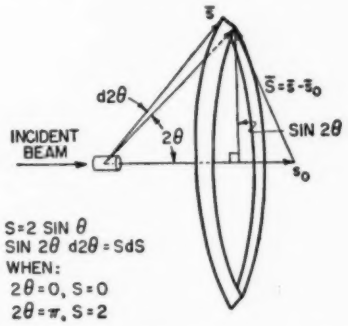


FIG. 1. The elastic scattering of neutrons by a powder and the differential solid angle involved in the integration to obtain the elastic cross section.

* Under the auspices of the U. S. Atomic Energy Commission.

† Guest Scientist at Brookhaven National Laboratory, Upton, New York.

[1] D. Kleinman, Phys. Rev. **81** (1951).

[2] Halpern, Hamermesh, and Johnson, Phys. Rev. **59**, 981 (1941).

[3] R. Weinstock, Phys. Rev. **65**, 1 (1944); Fermi, Sturm, and Sachs, Phys. Rev. **71**, 589 (1947).

[4] Weiss, Clark, Corliss, and Hastings, J. Appl. Phys. **23**, 1379 (1952).

[5] R. W. James, *The Optical Principles of the Diffraction of X-rays* (G. Bell and Sons, Ltd., London, 1954), Chapters IX and X.

$\bar{I}(S/\lambda)$ is the differential cross section per atom of the sample. $S=2\sin\theta$. N is the number of atoms per cm^2 of sample. The summations include all atoms in such a sample. The f's are the neutron scattering lengths, and γ_{jk} is the distance of separation between atoms j and k.

To find the cross section resulting from such scattering one must integrate over all solid angle. Referring to Fig. 1 the differential solid angle is $2\pi\sin 2\theta d2\theta$, and 2θ ranges from 0 to π.

$$\sigma(\Lambda)=2\pi\int_0^{\pi}\bar{I}(S/\Lambda)\sin 2\theta d2\theta. \tag{3}$$

Since $S=2\sin\theta$, $dS=\cos\theta d2\theta$, and $\sin 2\theta d2\theta=SdS$. The limits of S range from 0 when $2\theta=0$ to 2 when $2\theta=\pi$. Substituting in Eq. (3)

$$\sigma(\Lambda)=2\pi\int_0^{2}\bar{I}(S/\Lambda)SdS. \tag{4}$$

Let $x=S/\Lambda$. Then $x=0$ when $S=0$, and $x=2/\Lambda$ when $S=2$.

$$\sigma(\Lambda)=2\pi\Lambda^2\int_0^{x=2/\Lambda}\bar{I}(x)xdx. \tag{5}$$

If λ is the wavelength employed in a diffraction experiment and the differential cross section is plotted against $x=2\sin\theta/\lambda$, the cross section at the point Λ is simply $2\pi\Lambda^2$ times the first moment of the diffraction pattern.

The differential cross section can be expressed in terms of the cross section. Differentiating Eq. (5) with respect to Λ with due regard for the upper limit being a function of Λ, substituting using Eq. (5), and rearranging,

$$4\pi\bar{I}(2/\Lambda)=\sigma(\Lambda)-\frac{\Lambda}{2}\sigma'(\Lambda). \tag{6}$$

Equations (5) and (6) correlate the elastic cross-section plot and the diffraction pattern and allow derivation of one from the other. The same information, the interference function, is inherent in both. A cross section should be sensitive to crystalline particle size and strain content to the same degree as a diffraction pattern.

III. EXPERIMENT

To illustrate these conclusions, cutoffs exhibiting large strain and particle size broadening such as is found in carbonyl iron and exhibiting particle size broadening alone as in MgO have been transformed into diffraction patterns. These patterns were then compared to the x-ray diffraction patterns.

The transmission experiments employed a single (lead) crystal neutron spectrometer reflecting pile neutrons first filtered through beryllium. Incident and reflected beams were collimated with soller slits and the sample under study was placed directly in front of the BF_3 detector to eliminate loss by small angle scattering. Corrections for inelastic scattering from the lead monochromator and for background were then made. The thickness of the specimen was such that the transmission was approximately $1/e$. Corrections for extinction and multiple scattering were neglected. The x-ray patterns were obtained with a commercial recording diffractometer using filtered K_α radiation.

No correction for instrumental broadening effects either in the neutron or x-ray patterns was made. However, care was taken to make the instrumental resolution approximately the same for both cases. This was determined by comparing the x-ray and neutron line breadths from a well annealed iron sample. This breadth is indicated in the figures. Further the samples were deliberately chosen because their line breadths were considerably in excess of the instrumental breadths.

In Fig. 2(a) the dashed curve is the measured total cross section of S. F. grade carbonyl iron in the vicinity

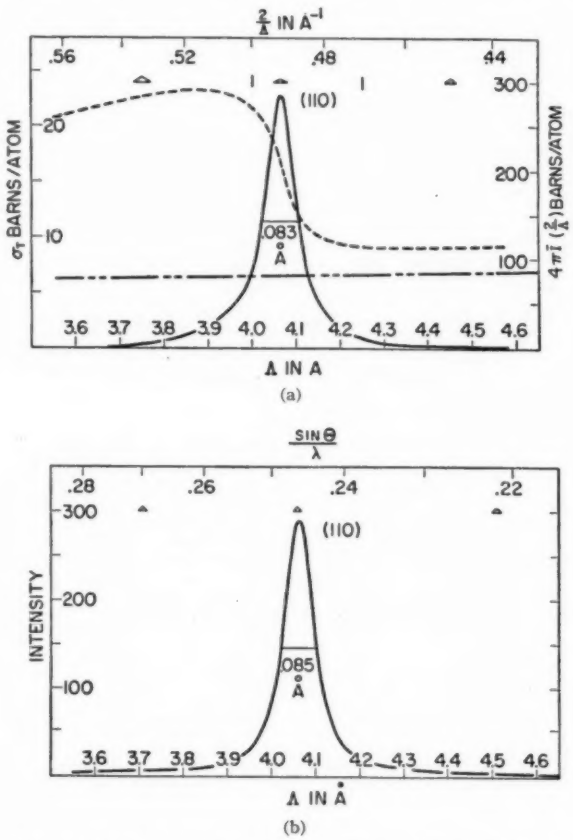


FIG. 2. (a) The total neutron cross section near the (110) cutoff of carbonyl iron (dashed), the inelastic and capture cross sections (broken line) and the derived (110) diffraction profile (solid curve) to be compared to that in (b). (Experimental points were too numerous for reproduction, but their scatter and the instrumental breadth are indicated by the heights and bases of the triangles). (b) The (110) x-ray diffraction profile of carbonyl iron, corrected for change in scattering and Lorentz-polarization factors.

of the (110) cutoff. The broken line was drawn through the origin and through the measured cross section beyond the cutoff for a specimen of well-annealed iron The difference between these curves was taken to be the elastic cross section. The solid curve is a diffraction profile derived using the transform in Eq. (6). This is to be compared to the pattern of Fig. 2(b) obtained from the x-ray profile of the (110) peak of the carbonyl iron corrected for change in scattering factor with $\sin\theta/\lambda$ and Lorentz-polarization factor. For neutrons the magnetic scattering is 8.6% of the total intensity in the (110) reflection[6] and no correction for form factor variation was made on the profile. Jaumot[7] has found a particle size of 70 A and a root mean square strain of 6×10^{-3} in the (110) direction for carbonyl iron.

The MgO was reduced from the carbonate to a particle size of 63.5 A. Warren,[8] Birks, and Friedman[9] have shown the size of MgO particles can be determined from the line broadening. The cross section of the MgO was found to be unexpectedly large. It is suspected that water was present in the form of impurity hydrates which did not decompose with the carbonate at 400°C. Large cross sections were also encountered in samples reduced at higher temperatures (1000°C) and in commercial MgO powders. In Fig. 3(a) the dashed curve is the total cross section of the MgO reduced at 400°C. The broken line was drawn according to the best estimate for capture and inelastic cross sections. The difference was taken to be the elastic cross section. Again the solid curve is a diffraction profile derived using Eq. (6) and is to be compared with the pattern of Fig. 3(b) obtained from the x-ray profile of the (200) peak of the MgO corrected for change in scattering and Lorentz-polarization factor.

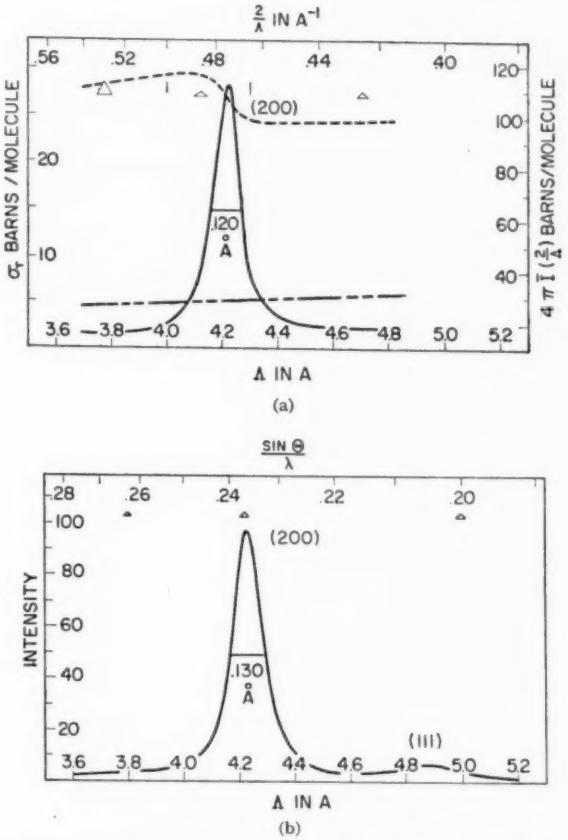


FIG. 3. (a) The total neutron cross section near the (200) cutoff of MgO (dashed), the inelastic and capture cross section (broken line) and the derived (200) diffraction profile (solid curve) to be compared to that in (b). (Experimental points were too numerous for reproduction, but their scatter and the instrumental breadth are indicated by the heights and bases of the triangles.) (b) The (200) x-ray diffraction profile of MgO corrected for change in scattering and Lorentz-polarization factors.

IV. DISCUSSION

There is good agreement between the derived and x-ray patterns for both carbonyl iron and MgO. The general line shapes agree, and the difference in half-widths of 2.4% and 8% are within the accuracy of the experiment. The transformations derived in this paper allow one to study diffraction profiles by transmission techniques. While only the profiles of single peaks were transformed, in principle the whole elastic cross section plot could be transformed to yield a diffraction pattern in absolute units. The theory applies to x-ray or neutron transmission although only the latter seems feasible. Finally, cross sections were shown to be sensitive to particle size and strain effects to the same degree as diffraction patterns, and comparisons between the two should be made by one or the other of the transformations given.

V. ACKNOWLEDGMENTS

The authors wish to thank Mr. R. F. Welter of the General Aniline and Film Corporation for the carbonyl iron powders.

[6] Shull, Wollan, and Koehler, Phys. Rev. **84** (1951).
[7] F. E. Jaumot, Jr., and L. Muldawer, Acta Metallurgica **2**, 513 (1954).
[8] B. E. Warren, J. Appl. Phys. **12**, 375 (1941).
[9] L. S. Birks and H. Friedman, J. Appl. Phys. **17**, 687 (1946).

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Radiation Induced Changes in the Electrical Resistivity of α Brass*

D. B. ROSENBLATT,† *Frankford Arsenal, Philadelphia, Pennsylvania,*
R. SMOLUCHOWSKI, *Carnegie Institute of Technology, Pittsburgh, Pennsylvania,*
AND
G. J. DIENES, *Brookhaven National Laboratory, Upton, New York*
(Received January 8, 1955)

Nucleon irradiation of metallic solid solutions may lower their resistivity by changing the degree of randomness (tendency to order or to cluster) or by a change of elastic moduli. Measurements of resistivity of brasses containing 10, 20, and 30% zinc have been performed at room temperature, at 80°K and at 4°K. The alloys were irradiated in a reactor at +50°C and at 80°K. It appears that the results can be best interpreted in terms of the formation of lattice defects upon low temperature exposure and increase of short-range order upon irradiation at +50°C. The change of residual resistivity is negative while the change of resistivity due to thermal scattering is positive for the samples irradiated at +50°C. The former depends upon irradiation flux and upon annealing time in a different manner from the latter. Short-range order appears to be produced also by proper heat treatment.

I. INTRODUCTION

NUCLEON irradiation of pure metals results in an increase of the electrical resistivity.[1–5] It is primarily the residual resistivity, ρ_0, which is increased by the introduction of crystalline defects although changes in the thermal part of the resistivity, ρ_T, have also been reported.[4] Decreases in resistivity have been observed in the order-disorder alloys AuCu and $AuCu_3$[6–8] during the early stages of irradiation of disordered and incompletely ordered specimens. This effect has been attributed to further ordering promoted by the irradiation although the details of the mechanism are by no means fully understood.[9]

Theoretical considerations[10] indicate that a decrease in resistivity may be observable in alloys in the absence of any ordering effects if thermal scattering is reduced due to a radiation induced increase in the elastic moduli.[11] Such a change may be observable in an alloy where the sensitivity of ρ_0 to the introduction of additional scattering centers may be small and the increase in ρ_0 may be outweighed by a decrease in ρ_T.

The above arguments provided the original stimulus for investigating the effects of reactor irradiation on α brass, an alloy in which there is no long-range order. Further motivation was provided by an observation by Billington and Kernohan[12] that relatively short reactor irradiation produces a small decrease in the resistivity of α brass. Consequently, it was decided to study this system in some detail. We may anticipate the results by stating that the original ideas were not confirmed, that is, the irradiation effects to be described cannot be attributed to change in ρ_T arising from changes in the elastic moduli. The currently available evidence points to alterations in short-range order as the principal, although probably not the only, effect.

II. APPARATUS

A number of alpha-brass polycrystalline wire samples were drawn to a diameter of about 0.025 in. and cut to lengths of approximately 8 cm, cleaned with acid and distilled water, and annealed for three hours at 350°C in a vacuum furnace and furnace cooled.[13] Copper and nickel wires for control experiments were prepared in a similar fashion except that the annealing was done at 450°C.

Two wires of identical composition were mounted on a Bakelite holder, one serving as a standard and the other as a sample. To each of these wires were soldered four short copper connectors, two to carry current and two for measuring the resultant potential drop. Provision was made for passing current through sample and standard in series and for carrying out the potential measurements through a cannon plug attached to the Bakelite holder. Separate holders were constructed for each material under investigation.

The resistance of a sample relative to a standard was determined before irradiation by plugging the holder into a circuit providing the dc current of a few amperes and allowing either of the potential drops to be switched to a type K Leeds and Northrup or

* Work done under contracts with the U. S. Atomic Energy Commission.
† Guest scientist at Brookhaven National Laboratory.
[1] For reviews see: J. C. Slater, J. Appl. Phys. **22**, 237 (1951); G. J. Dienes, Ann. Rev. Nucl. Sci. **2**, 187 (1953).
[2] Marx, Cooper, and Henderson, Phys. Rev. **88**, 106 (1952).
[3] R. R. Eggleston, Acta Metallurgica **1**, 679 (1953).
[4] D. Bowen and G. W. Rodeback, Acta Metallurgica **1**, 649 (1953).
[5] McReynolds, Augustyniak, McKeown, and Rosenblatt (to be published).
[6] T. H. Blewitt and R. Coltman, Phys. Rev. **85**, 384 (1952).
[7] L. G. Cook and R. L. Cushing, Acta Metallurgica **1**, 539 (1953).
[8] L. G. Cook and R. L. Cushing, Acta Metallurgica **1**, 549 (1953).
[9] T. H. Blewitt and R. R. Coltman, Acta Metallurgica **2**, 549 (1954).
[10] R. Smoluchowski, Phys. Rev. **91**, 245 (1953).
[11] G. J. Dienes, Phys. Rev. **86**, 228 (1952); **87**, 666 (1952); see also F. R. N. Nabarro, Phys. Rev. **87**, 665 (1952).
[12] D. S. Billington (private communication).
[13] Cooling rate of about 3°C/min.

Rubicon Model 2780 potentiometer. The ratio of the two potentiometer readings was found to vary less than 0.1% as the temperature of the bath in which the holder was immersed was changed from ambient to 80°K.

To determine the effect of reactor irradiation, samples were removed from their holders, but with their four copper connectors still attached, and were irradiated in the Brookhaven pile for stated periods either in a special hole maintained at +50°C by air cooling or in the liquid nitrogen irradiation facility.[5] Upon removal from the pile, samples were shielded and stored for approximately two weeks to allow the induced radioactivity to subside to a safe level. The samples were then remounted, care being exercised to heat only the tips of their connectors to avoid any annealing of the radiation effects in the sample, and the ratio of the resistances remeasured.

A modification of the apparatus was required for samples which were irradiated at 80°K. Since space limitations prevented irradiating a sample together with the connectors, a method had to be used in which the sample was measured before and after irradiation in a clip-type holder. This method is inherently less accurate and was found to suffer from roughly five times as great an experimental error.

III. EXPERIMENTAL RESULTS

(a) Effect of Reactor Irradiation of Alpha Brass for One Week at +50°C

Table I summarizes the results obtained from irradiating the samples in the +50°C facility of the B.N.L. pile for one week. This exposure corresponds to an integrated neutron flux of about 7×10^{17} nvt. The two nickel samples were cut from the same wire, whose purity was unknown, but was at least 99% Ni. The two 10% brass samples were also cut from a single piece of wire.

The percentage changes in resistivity for each sample are based upon means of 12 readings taken before and after irradiation. The deviations in these readings were such as to correspond to an error not greater than 0.1%. In order to evaluate the sources of error of a nonstatistical nature certain experiments were carried out, such as soldering and unsoldering, altering the temperature of measurement and subjecting the sample to oxidizing conditions similar to that in the reactor, etc. All these tests were also consistent with an over-all accuracy of 0.1%. It was found in an early experiment that when the samples were wrapped in Al foil appreciable surface discoloration occurred, probably due to radiation induced corrosion.[14] Discoloration was almost entirely eliminated when the samples were irradiated in an Al capsule and especially in a quartz capsule.

TABLE I. Irradiation effect: one week at +50°C. Figures give percentage change in resistivity due to the irradiation as obtained by measuring relative resistances before and after irradiation. The temperatures shown are the temperatures of measurement. (Concerning the accuracy of 4°K measurements, see text.)

Sample	$\Delta R\%$ 4°K	$\Delta R\%$ 80°K	$\Delta R\%$ room temp.	
Cu	...	+1.1	+0.2	Pure Metal
Ni	...	+0.8	+0.2	
Ni	...	+0.9	+0.2	
10% Zn	...	−1.5	−0.7	
10% Zn	−1.9	−1.4	−0.7	α Brass
20% Zn	−4.4	−1.9	−0.5	
30% Zn	−0.8	−2.1	+0.2	

Table II. Absolute resistivity changes induced by one-week irradiation at +50°C. Resistivities are given in micro-ohm centimeters.

	$\Delta\rho_0$[a]	$\Delta\rho_T$ 300°K	$\Delta\rho_T$ 80°K	ρ_0 (4°K)	ρ_T 300°K	ρ_T 80°K	ρ 300°K	ρ 80°K
10% Zn	−0.034	+0.006 (±0.008)	+0.003 (±0.004)	1.8	2.2	0.3	4.0	2.1
20% Zn	−0.132	+0.10	+0.07	3.0	2.4	0.35	5.4	3.35
30% Zn	−0.028	+0.04	−0.05	3.48	2.82	0.43	6.3	3.91

[a] Concerning uncertainty in $\Delta\rho_0$, because of error in the 4°K measurements, see text.

The values for the changes in resistivity measured at 4°K are somewhat unreliable since the degree of error is unknown. This is due to the fact that the resistance ratios were measured only after irradiation and compared with an estimate of the resistance ratios at 4°K before irradiation. Annealing results to be discussed later are consistent with the data recorded in Table I.

The primary reason for the inclusion of pure metal control samples was to provide a check on the integrated neutron flux. The percentage increase in resistivity of the Cu samples agrees well with the values to be expected for normal pile operation for a duration of one week if the annealing at 50°C is taken into account.[5] It may be of interest to note that, since the resistivity of Ni is much higher than that of Cu, the change obtained in Ni implies a larger $\Delta\rho$ for this metal. This is consistent with the cyclotron studies of Marx, Cooper, and Henderson.[2]

In order to separate the radiation induced changes in resistivity into $\Delta\rho_0$ and $\Delta\rho_T$, a knowledge of the absolute resistivity of alpha brass was required. Absolute resistivities for 10% Zn and 30% Zn brass were measured[15] at 4°K, 80°K, and at room temperature. These were found to be very close to values obtained by extrapolation of data reported in the literature.[16] The absolute resistivities used in all calculations are given in Table II. These values have been used for the conversion of the observed percentage changes into $\Delta\rho_0$ and $\Delta\rho_T$ as shown in Table II.

If the decrease in resistivity observed in alpha brass after irradiation for one week at 50°C is the result of atomic rearrangement, then annealing for a sufficient time at a sufficiently elevated temperature should

[14] W. Primak and L. H. Fuchs, Phys. Today **7**, 15 (1954).

[15] We are indebted to Professor S. Friedberg for performing for us preliminary measurements on a 30% Zn brass.

[16] C. S. Smith, Trans. Am. Inst. Mining Met. Engrs. **89**, 84 (1930); **117**, 225 (1935).

TABLE III. Annealing data for 10% Zn sample after irradiation at +50°C for one week (annealed for 22 hr at 160°C in a vacuum furnace).

Temp. of measurement	4°K	80°K	Room
Irradiation produced change	−1.9%	−1.4%	−0.7%
Annealing after irradiation	+2.5%	+1.6%	+0.9%
Overshooting	0.6%	0.3%	0.2%

result in complete return of the resistivity to the value measured before irradiation. Unfortunately, annealing at high temperature could not be carried out, since soft solder was used for the connectors. However, as is shown in Table III, recovery appears to be complete for an anneal of 22 hr at +160°C in a vacuum furnace.[17]

An attempt was made to determine the temperature region in which the recovery of the resistance first sets in. In the case of the 10% Zn sample rapid recovery of the resistivity does not occur until +170°C, and at +180°C 20% of the resistivity decrease has annealed out in one hour. The recovery in the case of the 20% Zn sample appears to set in at a slightly lower temperature. These results show that a recoverable radiation effect has been induced by reactor exposure.

Upon examination of Table I it is evident that when measurements are carried out at 80°K, an irradiation for one week at +50°C leads to the expected increase in resistivity of pure metals, but to a decrease of 1% or 2% for each of the alpha-brass samples. The changes observed at room temperature are somewhat smaller, particularly for the 30% Zn sample. Thus it appears almost certain that the primary effect of the radiation is to reduce the residual resistivity; any change in the Debye temperature must be relatively small since such a change would affect the resistivity more markedly at room temperature than at 80°K. Moreover the change in elastic constants referred to in the introduction should result in negative values for $\Delta\rho_T(300°K)$ whereas Table II indicates that $\Delta\rho_T(300°K)$ is invariably positive. These considerations appear sufficient to rule out a change in elastic constants as a principal cause of the observed resistivity decrease produced by the one-week irradiation in alpha brass.

(b) Reactor Irradiation at 80°K

Since the major change produced by the irradiation at +50°C for one week was a decrease in the residual resistivity, it seemed natural to inquire whether the effect might not be related to ordering or to clustering. In order to do so additional irradiations and measurements were carried out at 80°K where the low mobility of defects formed by the radiation would prevent any radiation induced ordering. If the observed decrease in resistivity at 80°K due to an irradiation at +50°C is produced by the decrease in ρ_T associated with elastic constant changes, this drop in resistivity should be more noticeable the greater the concentration of defects, and consequently, the resistivity drop should be more pronounced if irradiation and subsequent storage takes place at a temperature low enough to "freeze in" most defects.

The results of irradiation at 80°K for one month (about 3×10^{18} nvt) are shown in Table IV. An increase in resistivity is observed in every case when the resistivity is measured at liquid nitrogen temperature. It is interesting to calculate the increase in the absolute resistivity of α brass due to this low temperature irradiation. The data are given in Table IV. It should be noted that the absolute changes in resistivity for the three kinds of α brass lie between the values observed elsewhere[5] for Cu and Al. The value for 10% Zn brass is not much greater than that for Cu, the value for 30% Zn brass being almost as great as that for aluminum. These results suggest that the change in ρ increases with zinc content because a higher proportion of displaced atoms are zinc ions, which, being more highly charged, are more efficient scattering centers for the conduction electrons.

The annealing out of the radiation effect which was induced by irradiation at 80°K is quite similar to that in pure metals. Two examples are given in Fig. 1 where the resistivity increment remaining in the sample is plotted against the annealing time. All measurements were made at 80°K. It is seen from Fig. 1 that a 20°C anneal removes completely the low temperature radiation effect in the 10% Zn sample. The curve for the 20% Zn sample indicates that annealing proceeds at quite a fast rate even at −32°C. The main point in the present connection is that radiation effects introduced at low temperature are completely removed by a long time anneal in the 25–50°C range. The low-temperature radiation effect in α brass is, therefore, very similar to that in pure metals. All the evidence indicates that essentially no defects are preserved during a 50°C irradiation, at least as measured by resistivity changes. Vacancies and interstitials, therefore, cannot be directly responsible for the resistivity decrease resulting from a 50°C irradiation.

(c) Change in Resistivity of Unirradiated α-Brass Samples Induced by Means of Prolonged Heating

The results of the experiments at 50°C described so far are most easily explained by a hypothesis of radia-

TABLE IV. Increase in absolute resistivity due to irradiation for one month at 80°K, measured at 80°K.*

Material	$\Delta R/R$	ρ Micro-ohm cm	$\Delta\rho$ Micro-ohm cm
Cu	16.%	0.23	0.0368
α Brass, 10% Zn	2.7%	2.1	0.0570
α Brass, 20% Zn	3.2%	3.4	0.109
α Brass, 30% Zn	3.1%	4.0	0.124
Al	25.%	0.64	0.16

* Note: Data for Cu and Al taken from reference 5.

[17] Further detailed experiments will be required to define unambiguously the state of short-range order in the annealed specimens.

tion induced ordering in α brass. A principal objection to such a hypothesis would be the lack of evidence for an ordered state in α brass induced by agents other than radiation. An exception is the experiment of Masumoto, Saito, and Sugihara[18] who found a specific heat anomaly in α brass which they ascribed to ordering.[19] Using their data the temperature most suitable for a thermal ordering experiment was estimated to be +190°C. The duration of this heat treatment was arbitrarily selected as six weeks.

Table V summarizes the results obtained by heat treating 20% and 30% Zn α-brass samples. These samples were constructed with silver soldered connectors (instead of the usual soft solder) so that they could withstand the heat treatment. Before commencing the experiment, the 20% Zn and 30% Zn samples were thoroughly annealed by heating at 370°C for 1½ hours in vacuum and cooled by shutting off the heating current and opening the furnace.[20] The resistivities

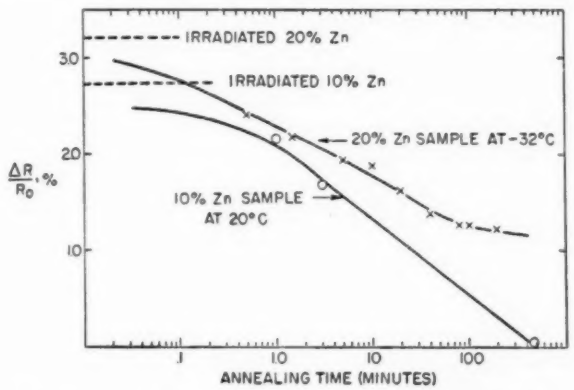


FIG. 1. Annealing of resistivity changes produced in α-brass samples by irradiation at 80°K. The resistivity was measured at 80°K and the annealing temperatures are indicated on the graphs. R_0 refers to the unirradiated sample.

were measured at 80°K and at room temperature. The samples were then heated at +190°C for six weeks in vacuum. Upon remeasurement, the resistivities of both samples were found to have suffered a significant decrease. It should be noted that the percentage decreases as well as the $\Delta\rho$ values produced by the heat treatment are greatest in magnitude the lower the temperature of measurement. This favors an interpretation based on thermally induced ordering. It should be noted that the amount of change in the absolute resistivity produced by the heat treatment is independent of zinc content, or almost so, within experimental error. It should be further noted that the magnitude of the resistivity changes are comparable

[18] Masumoto, Saito, and Sugihara, Sci. Repts. Research Inst. Tôhoku Univ. A4, 481 (1952).

[19] In a private communication, Dr. P. Gordon of Illinois Institute of Technology informed us that the kinetics and size of heat effects obtained in his recent calorimetric studies on 70: 30 brass are suggestive of short-range ordering.

[20] Cooling rate of about 11°C/min.

TABLE V. Resistivity changes produced in unirradiated 20% and 30% Zn samples of alpha brass by heating at +190°C for a period of six weeks.

Temp. of measurement	% Zn	% change in resistivity	Absolute resistivity $\rho\ 10^{-6}$ ohm-cm	Change in absolute resistivity, $\Delta\rho\ 10^{-6}$ ohm-cm
80°K	20%	−1.11±0.03	3.35	−0.037
	30%	−0.92±0.03	4.13	−0.038
Room	20%	−0.30±0.03	5.4	−0.016
	30%	−0.33±0.05	6.30	−0.021

to those produced by an irradiation of one week at +50°C.

If the effect of heating for six weeks at +190°C is caused by an actual change within the lattice of alpha brass and not some spurious effect such as loss of zinc during the treatment, it should be possible to restore the resistivities of the samples to their original values by thorough annealing. Data are available only for the 20% Zn sample which were obtained by subjecting the sample to the same annealing procedure (1½ hours at 370°C) as was used at the start of this experiment. The original resistivity was restored. As a matter of fact, the final resistivity was found to be slightly greater than the initial resistivity. Measurements at 4°K were also made before and after this anneal. The data in Table VI show that the heat treatment (6 weeks at +190°C) has lowered the residual resistivity, and permits the separation of the change in absolute resistivity caused by the heat treatment into the change in ρ_0 and in ρ_T. Comparison of the changes $\Delta\rho_0$ and $\Delta\rho_T$ induced by heat treatment and by irradiation at +50°C for one week (see Table VI) shows that in both cases $\Delta\rho_0$ is negative and $\Delta\rho_T$ positive and that $\Delta\rho_0$, $\Delta\rho_{T(80°K)}$ and $\Delta\rho_{T\text{room}}$ are each about doubled in going from the heat treated to the irradiated sample. It appears that the same type of lattice rearrangement can be produced either by heat treatment or irradiation.

The changes observed for the heat treated samples are suggestive of ordering since $\Delta\rho_0$ is found to be negative. In order to check this further a single crystal of α brass (30% Zn) was heat treated the same way as the wire sample used for the resistivity studies. Keating[21] carried out a neutron diffraction study on this sample in an attempt to verify the ordering

TABLE VI. Comparison of changes in $\Delta\rho_0$ and $\Delta\rho_T$ produced by heat treatment and by irradiation one week at +50°C for a 20% Zn sample.

Temp. of measurement	Changes due to irradiation at +50°C for 1 week 10^{-6} ohm-cm	Changes due to 6 week anneal at +190°C 10^{-6} ohm-cm
4°K	$\Delta\rho_0 = -0.132$	$\Delta\rho_0 = -0.069$
80°K	$\Delta\rho_T = +0.069$	$\Delta\rho_T = +0.032$
Room	$\Delta\rho_T = +0.105$	$\Delta\rho_T = +0.053$

[21] D. Keating, Acta Metallurgica (to be published).

TABLE VII. Resistivity changes produced by a one-month irradiation at +50°C of 20% and 30% Zn samples.

Zn %	Temp. of measurement	$\Delta R/R$%	ρ 10^{-6} ohm-cm	$\Delta\rho$ 10^{-6} ohm-cm
20%	80°K	+0.16±0.1	3.35	+0.00536
	room	+1.9 ±0.1	5.4	+0.103
	4°K	−4.4*	3.0	−0.132*
30%	80°K	+0.38±0.1	4.13	+0.0157
	room	+2.2 ±0.1	6.3	+0.139

* This measurement is of doubtful accuracy.

hypothesis. He found that long-range order was absent ($S<0.0365$) and that no additional diffuse reflection was observable. However, a careful analysis showed that a considerable amount of short-range order could be present and not detectable either with neutrons or x-rays. Thus, while the ordering hypothesis could not be verified, the diffraction data were not in contradiction to such a hypothesis.

(D) Reactor Irradiation of α Brass for One Month at +50°C

Irradiation experiments were also carried out on α brass for one month at +50°C (about 3×10^{18} nvt). The resistivity data are shown in Table VII. A resistivity increase, rather than decrease, was found when measured at 80°K or at room temperature. A doubtful measurement at 4°K for the 20% Zn sample indicates that $\Delta\rho_0$ is still negative and of about the same order-of-magnitude as obtained after the one-week irradiations. The very much smaller $\Delta\rho$ observed at 80°K as compared to the room temperature measurement points quite conclusively in the same direction. Thus, the reversal in sign after prolonged irradiation is almost certainly due to a large positive $\Delta\rho_T$ relative to a smaller change (if any) in the negative $\Delta\rho_0$.

In an attempt to verify the foregoing conclusions, some annealing experiments were carried out at 175°C with measurements taken at 80°K and at room temperature. The data are shown in Fig. 2. An increase in resistivity is observed in the early stages of the annealing with this increase particularly prominent at 80°K. At room temperature, where ρ_T plays a far more prominent role, a reversal is shown as the annealing proceeds. These data confirm, therefore, the interpretation that a rather large negative $\Delta\rho_0$ must have been produced by the irradiation. The data also indicate that $\Delta\rho_0$ and $\Delta\rho_T$ do not anneal out together.

IV. DISCUSSION

The two most important experimental results emerging from this study are:

(1) With certain choices of integrated fast neutron flux and temperature of irradiation, it is possible to reduce the resistivity of alpha brass by 1 to 2%.

(2) Prolonged heating at +190°C will produce a decrease in the resistivity of alpha brass of about 1%.

The fact that the changes observed are largest at the low temperature of measurement appears to rule out the change in elastic constants as an important factor. This is further substantiated by the fact that irradiation at 80°K produces an increase in resistivity in every way analogous to that produced in pure metals.

The most likely cause of the decrease in resistivity observed during the +50°C irradiation and that due to the +190°C heat treatment is short-range ordering. The hypothesis of short-range order is consistent with the fact that, regardless of whether heat treatment or irradiation is used to cause the decrease in resistivity, the change can always be broken up into a $\Delta\rho_0$ which is negative and a $\Delta\rho_T$ which is positive. Short irradiations result in a large negative $\Delta\rho_0$ and a small positive $\Delta\rho_T$, thereby resulting in a negative $\Delta\rho$. Long irradiations produce a large positive change in ρ_T and therefore result in a positive $\Delta\rho$ when measured at temperatures where ρ_T is important. There is some indication that the decrease in $\Delta\rho_0$ saturates as a function of irradiation which is consistent with the idea that an equilibrium short-range order is being approached during irradiation.

The various annealing experiments are important in demonstrating that the resistivity changes are reversible. They are, however, somewhat difficult to interpret. For example, the recovery of the resistivity in α brass, previously irradiated for one week at +50°C, in the 160°C range may seem somewhat puzzling when one considers that annealed brass undergoes a resistivity decrease when heated at 190°C for prolonged times. These results, however, are probably not contradictory because of the difference in temperature and consequent different rates of ordering. Perhaps the most curious results are the annealing data on samples irradiated for one month at 50°C. In this case $\Delta\rho_0$ and $\Delta\rho_T$ do not recover at the same time. The negative $\Delta\rho_0$ apparently anneals out first. This seems to indicate that, although the negative $\Delta\rho_0$ and the positive $\Delta\rho_T$ are produced concurrently during irradiation, at least

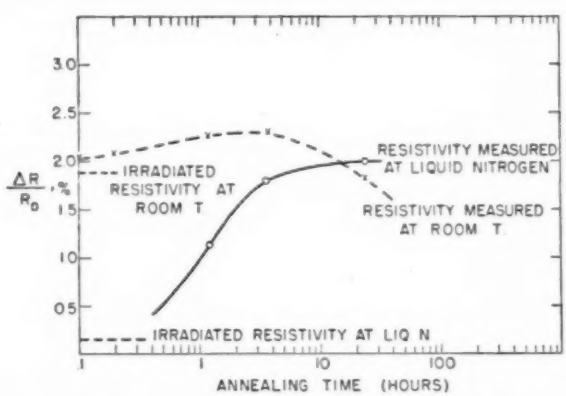


FIG. 2. Annealing of the change in resistivity produced by one month irradiation of a 20% Zn sample at 50°C. Annealing temperature: 175°C. Temperature of measurement as indicated on the graphs. R_0 refers to the unirradiated sample.

part of the $\Delta\rho_T$ may not be ascribable to change in short-range order.

V. SUMMARY

The results may be summarized briefly as follows: Alpha brass possesses states of various degrees of short-range order at different temperatures. The degree of short-range order may be increased by a suitable reactor irradiation or heat treatment. An increase in order is associated with a negative $\Delta\rho_0$ and a positive $\Delta\rho_T$. The decrease in $\Delta\rho_0$ probably saturates as a function of irradiation while $\Delta\rho_T$ apparently continues to increase. After a one-month reactor irradiation at 50°C $\Delta\rho_0$ and $\Delta\rho_T$ do not anneal out the same way. Consequently, part of the increase in $\Delta\rho_T$ probably cannot be attributed to the alteration of short-range order.

ACKNOWLEDGMENTS

We are indebted to A. W. McReynolds, Marilyn McKeown, and W. M. Augustyniak for advice concerning general techniques, Fred Iseli for carrying out the six week's heat treatment of samples, J. K. Dominy for assistance in setting up the experiments at liquid helium temperature, H. I. Fusfeld and A. Nowick for helpful discussions, J. J. Floyd for carrying out the irradiations at +50°C, and Jerome Fleeman for the irradiations at liquid nitrogen temperature.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 8 AUGUST, 1955

Dependence of the Coercive Force on the Density of Some Iron Oxide Powders*

A. H. Morrish and S. P. Yu
Department of Electrical Engineering, University of Minnesota, Minneapolis, Minnesota
(Received January 24, 1955)

An experimental investigation has been made of the dependence of the coercive force on the density of some magnetite and gamma ferric oxide powders consisting of different sized and shaped particles. Various aspects of the data are discussed. In particular, an interpretation based on a calculation of the critical single domain size of iron oxide particles with a prolate spheroidal shape is consistent with the experimental results.

INTRODUCTION

THE study of the properties of small particles of magnetic materials, particularly of iron, cobalt, nickel, and a few alloys, has received considerable attention in the last decade, and several important advances in theory and in experiment have been made.[1] Evidence has been put forward for the existence of single domain structure for particles below a certain size. This size depends on the shape and internal strains of the particles, as well as on the material itself. To the extent to which experiment could be compared with theory, there exists rough agreement.

This paper reports on a study of two iron oxide powders, magnetite (Fe_3O_4) and gamma ferric oxide (γ-Fe_2O_3). Studies of such powders, of interest in themselves, also are of value in the understanding of the properties of that important class of materials known as the ferrites. In fact, magnetite, from which the ferrites are derived, may properly be called ferrous ferrite. Such substances crystallize in the *spinel* structure. Basically the structure consists of a closely packed lattice of oxygen ions, with the cations occupying certain interstitial positions. These are usually called *A* sites and *B* sites, the former positions being surrounded tetrahedrally by four oxygen ions, the latter by six oxygen ions placed octahedrally. With magnetite, in a unit cell containing thirty-two oxygen anions, eight *A* sites and eight *B* sites are occupied by ferric ions, while eight *B* sites are occupied by ferrous ions.[2] Gamma ferric oxide is a special case, where on the average 2-2/3 cation type *B* sites per unit cell are empty.[3] The other well-known iron oxide, α-Fe_2O_3, does not have the spinel structure.

Early work on the iron oxides, summarized by Koenigsberger,[4] has primarily investigated the variation of coercive force as a function of particle size. However, the smallest particles studied were still apparently much larger than the critical size. Welo and Baudisch[5] have listed methods of producing some iron oxide powders chemically. Recently, Osmond[6,7] has put forward an interpretation of the magnetic properties of some powders used as the active coating of sound-recording tapes.

In our experiment, measurements of the coercive force as a function of the density of some powder

* This work has been supported financially by the Mines Experimental Station of the University of Minnesota as part of their taconite program.

[1] See, for example, review articles by C. Kittel, Revs. Modern Phys. **21**, 541 (1949); and L. Weil, J. phys. radium **12**, 437 (1951).

[2] E. J. W. Verwey and E. L. Heilmann, J. Chem. Phys. **15**, 174 (1947).

[3] L. Néel, Ann. phys. **3**, 1937 (1948).

[4] J. G. Koenigsberger, Phil. Mag. **38**, 640 (1947).

[5] A. Welo and O. Baudisch, Chem. Rev. **15**, 45 (1934).

[6] W. P. Osmond, Proc. Phys. Soc. (London) **B65**, 122 (1952).

[7] W. P. Osmond, Proc. Phys. Soc. (London) **B66**, 265 (1953).

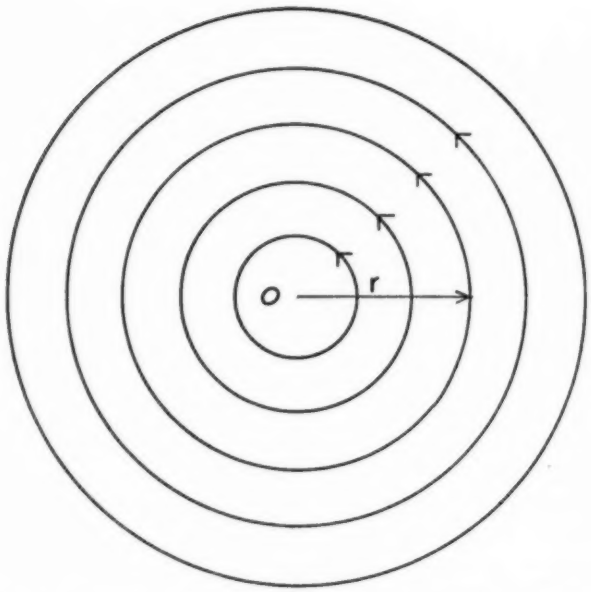


FIG. 1. Illustration of flux closure produced by a ring configuration in the equatorial plane of the prolate spheroidal particle.

samples are made. These samples contain various sized and shaped particles of Fe_3O_4 and γ-Fe_2O_3, close to the critical size limit. Previously, Osmond[6] had calculated the critical size of magnetite and gamma ferric oxide particles using as a model the crystal structure of iron. This is obviously not correct. We therefore begin by outlining, in the next section, a calculation of the critical size for these materials based on the spinel crystal structure. Other pertinent formulas relating to single domain particles are collected at the end of this section. We proceed by presenting the experimental results, and the paper is concluded with a discussion of these results.

CRITICAL SINGLE DOMAIN SIZE

First, we propose to calculate the critical single domain size for a magnetite particle, with the shape of a prolate spheroid. The calculation is similar in principle to that carried out by Néel,[8] Kittel,[9] and Stoner and Wohlfarth[10] for the magnetic metals, but differs in details because of the greater complexity of the spinel lattice. For the present, the effect of the magnetocrystalline anisotropy will be neglected. The increase in exchange energy when the spins are arranged in a circular configuration as shown in Fig. 1 is first determined. With this arrangement there are no magnetic poles, and the magnetostatic energy is zero. This exchange energy is compared with the magnetostatic energy that would be produced by the particle if it had a single domain configuration. The particle size, below which the energy of the single domain arrangement is less than the energy of a closed flux structure, is the critical size.

We proceed to calculate the exchange energy for the circular spin arrangement of Fig. 1. Barring other factors, the spins of the ions would tend to be aligned parallel to one of the easy directions of magnetization. For magnetite, at room temperature, these easy directions are along the body diagonals. Since two of these directions lie in the (110) plane, changes in angle between the spins of successive ions would occur in a direction perpendicular to this plane. Figure 2 shows a section through a (001) plane, containing oxygen ions and cations in octahedral positions. The large circles represent the oxygen ions, and the smaller, hatched circles the cations in B sites. The other circles refer to A site cations just above (M) or below (N) the plane. The dotted lines that form a square outline one unit cell of the spinel. We consider then the exchange energy along the horizontal direction (a_1).

A complete solution would require that we take account of the A–A, B–B, and A–B interactions. However, we further simplify the problem by considering only the predominant A–B interaction. We denote by $\pm J_{AB}$ the A–B exchange energy per atom pair for parallel and antiparallel spin alignment, respectively. Further, because the separation of the A and B cations by the larger oxygen ions is large compared to the average radius of the iron ions, we consider the four or five effective electrons of the ferrous and ferric ions to act as though they were localized at the center of an A or B site. It will not be necessary to consider the detailed nature of the A–B superexchange interaction. We assume that a reasonable approximation for J_{AB} is given by

$$J_{AB}=I_0^2k\theta/I^2_{00}p, \tag{1}$$

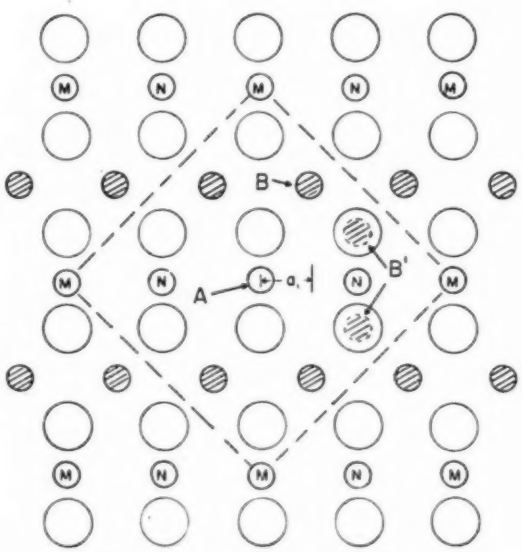


FIG. 2. The ions lying in a (001) plane of a spinel. The large circles represent oxygen ions, and the hatched circles the B or octahedral sites. The other small circles represent A sites, either slightly above (M) or below (N) the plane of the diagram.

[8] L. Néel, Compt. rend. **224**, 1488 (1947).
[9] C. Kittel, Phys. Rev. **70**, 965 (1946).
[10] E. C. Stoner and E. P. Wohlfarth, Trans. Roy. Soc. (London) **A240**, 599 (1948).

where k is Boltzmann's constant, θ is the Curie temperature ($=848°$K), I_0 and I_{00} are the saturation intensities at room temperature and at absolute zero, respectively. The number of nearest neighbors, p, for A or B sites per molecule, is equal to 12, since each A site has 12 B nearest neighbors, while each of the two B sites has 6 nearest A neighbors.

Considering only nearest neighbor interactions, the exchange energy per unit volume is

$$E_{ex.}=-\tfrac{1}{2}n\sum_{p}J_{AB}\cos\epsilon, \tag{2}$$

where n is the number of atoms in a unit volume that occupy a particular site, the unit volume being taken to be one molecule of Fe_3O_4. The angle between the spin directions of nearest neighbors is denoted by ϵ. For a ferric ion in an A site the exchange energy is

$$E_A=-\tfrac{1}{2}n_AJ_{AB}[2\cos0°+6\cos\epsilon+4\cos2\epsilon], \tag{3}$$

since $2B$ ions are at the same distance from the (110) plane as A, $6B$ ions lie in a (110) plane a distance a_1 either ahead or behind A, (site B in Fig. 2), and $4B$ ions lie in a (110) plane a distance $2a_1$ ahead or behind A, (sites B' lying above the plane of the diagram, indicated by the dotted circles in Fig. 2). Verification of this result is most easily seen by referring to a three-dimensional model of the spinel lattice.

On expanding the cosine terms of Eq. (3), collecting, and considering only the variable part, we get

$$E'_A=\frac{11}{2}n_AJ_{AB}a_1^2\left(\frac{d\phi}{dn}\right)^2, \tag{4}$$

where ϕ is the angle a spin vector makes with the z-axis, and n is measured along the horizontal direction.

The variable part of the exchange energy due to the interaction of A sites on the $2B$ sites is of course equal to E'_A. A check on the result is obtained by noting that for one of the B sites $E_{B(1)}=-\tfrac{1}{2}n_{B(1)}J_{BA}[6\cos\epsilon]$, while for the other B site $E_{B(2)}=-\tfrac{1}{2}n_{B(2)}J_{BA}[2\cos0°+4\cos2\epsilon]$. Therefore, the total variable exchange energy per molecule of Fe_3O_4 is:

$$E'_{mol.}=\frac{11}{12}\frac{I_0^2}{I_{00}^2}k\theta a_1^2\left(\frac{d\phi}{dn}\right)^2. \tag{5}$$

We proceed to calculate the exchange energy of the molecules in a ring of radius r, as indicated in Fig. 1. The total change in angle of the spins in going around the ring is 2π, and the circumference of the ring is $3a_1N$, where N is the number of molecules, i.e., $d\phi/dn=2\pi/3a_1N$. The number of molecules in such a ring is $2\pi r/s$, where s is the cell constant of one molecule of Fe_3O_4, and equal to 4.2×10^{-8} cm. Therefore $d\phi/dn=s/3a_1r$.

The exchange energy per ring is then $E'_{ring}=E'_{mol.}\cdot2\pi r/s$.

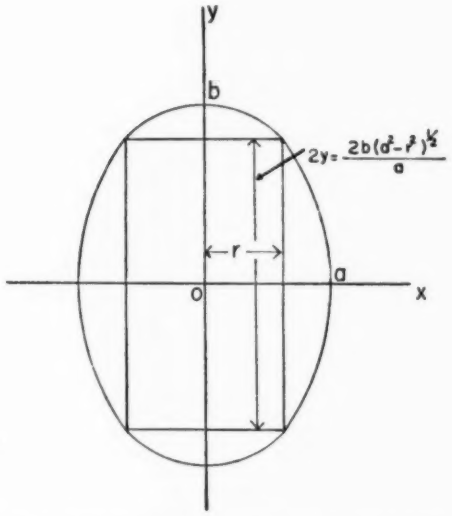


FIG. 3. Prolate spheroid considered as composed of shells of circular cylinders.

Let us consider a prolate spheroid, with major axis b and minor axis a. Now the number of ring in a cylindrical shell as shown in Fig. 3 is $2y/s$ where $y=b(a^2-r^2)^{\frac{1}{2}}/a$; therefore

$$E'_{cyl.}=E'_{ring}\cdot\frac{2b(a^2-r^2)^{\frac{1}{2}}}{as},$$

and the exchange energy for the prolate spheroid is

$$E'_{sph.}=\int_s^a E'_{cyl.}\cdot\frac{dr}{s}$$

$$=\frac{11}{12}\cdot\frac{I_0^2}{I_{00}^2}\cdot\frac{4\pi}{9}\cdot\frac{k\theta b}{as}\int_s^a\frac{(a^2-r^2)^{\frac{1}{2}}}{r}dr.$$

For $a\gg s$ this gives finally,

$$E'_{ex.}=\frac{11}{12}\cdot\frac{I_0^2}{I_{00}^2}\cdot\frac{4\pi}{9}\cdot\frac{k\theta b}{s}\cdot\ln\left(\frac{2a}{s}-1\right). \tag{6}$$

On the other hand, the magnetostatic energy for a saturated single domain configuration is $E_{mag.}=\frac{1}{2}I_0^2D\times[(4\pi/3)a^2b]$ for a similar prolate spheroid. Here D is the demagnetization factor of the particle.

The critical size is found by equating (6) to the magnetostatic energy. We get

$$\frac{a^2}{\ln476a\times10^{-7}-1}=\frac{6.18\times10^{-12}}{D}. \tag{7}$$

The length ($2b$) as a function of a/b for magnetite, as obtained from this equation is shown in Fig. 4.

The calculation of the critical size for γ-Fe_2O_3 proceeds as for magnetite, except that one must take account of the fact that only $13\frac{1}{3}/16$ of the B sites are filled. It turns out that the critical size for γ-Fe_2O_3 is

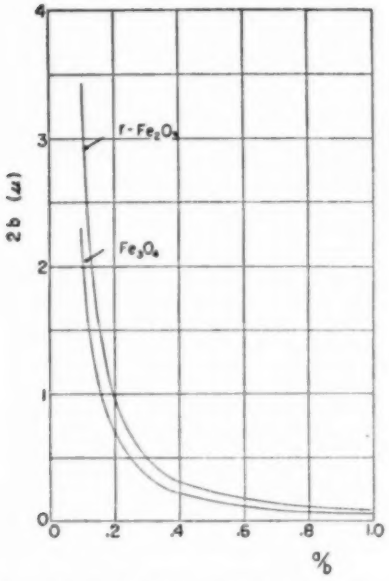


FIG. 4. The critical single domain size of Fe_3O_4 and γ-Fe_2O_3. The length of the major axis of the prolate spheroid is shown as a function of minor to major axis ratio.

somewhat larger than for Fe_3O_4. The results of the calculation are also shown in Fig. 4.

It is permissible to neglect the crystalline anisotropy only if the particle width ($2a$) is less than $\sqrt{2}$ times the thickness of a Bloch wall for the material. The approximate thickness of a Bloch wall may be calculated by minimizing the sum of the exchange and anisotropy energy of the wall. The result of such a calculation gives a thickness of 1.38×10^{-5} cm for Fe_3O_4. It follows that the magnetic particle width must be at least 1.95×10^{-5} cm before complete multidomain structure can exist. It turns out, then, that neglect of the magnetocrystalline anisotropy is justified except for particles with $b/a>7$. Inclusion of this factor for these particles would increase the critical size somewhat. For example, for a particle with $b/a=10$, the length $2b$ is increased by approximately 40% whereas for $b/a=5$ the length is increased by only 3%.

We have considered other possibilities for a flux closure arrangement. Moments in a simple elliptical configuration, lying in a plane parallel to the long axis of the spheroid, are not permitted because of the large amount of anisotropy energy. An arrangement, with a Bloch wall along the particles' b axis, together with a circular type arrangement at the ends, results in a particle size for the extreme case ($b/a=10$) that is slightly larger than for the pure circular scheme. Further, a calculation for a spherical particle, assuming two mutually perpendicular Bloch walls, gives a critical size much smaller than the Bloch wall thickness, and should therefore be disregarded.

Because of the nature of the approximations, the calculated critical sizes may be incorrect by a factor of about 2. These critical sizes likely represent, however, more of a lower than a higher limit, since the estimate of the exchange energy made is probably too small.

An alternative criterion that could be used to determine whether or not a powder consists of single domain particles, is the comparison of the coercive force with that predicted by theory for single domain particles. According to Stoner and Wohlfarth,[10] the coercive force, due only to shape anisotropy, of a powder whose single domain particles are randomly orientated, and assuming no interparticle interactions, is 0.479 $(D_b-D_a)I_0$, where D_a and D_b are the demagnetization factors along the particles minor and major axis, respectively. For example, if $b/a=5$, $I_0=480$ for Fe_3O_4 and 400 for γ-Fe_2O_3, one obtains 1170 and 976 oersteds, respectively, as the coercive force for the two powders. On the other hand,[8] if the single domain particles are spherical in shape, so that the crystalline anisotropy is the governing factor, and again assuming random orientation and no particle interactions, $H_c=0.64|K|/I_0$. For magnetite this gives a value of $\sim$145 oersteds.

EXPERIMENT

The coercive force, H_c, measured in these experiments, was the value of the applied field, in oersteds, for which the magnetization of the sample was zero. The apparatus consisted of a long air cored solenoid, and a small pickup coil, mounted rigidly at the center of the solenoid, connected to a galvanometer. The reverse field for which motions of the magnetic sample, suspended in the pickup coil produced no galvanometer deflection, was then the coercive force. Fields of up to 3000 oersteds, required to saturate the sample, were obtained by discharging a bank of condensers through the solenoid, oscillations being prevented by a thyratron in the circuit. The time of the discharge was about 1/60 seconds.

One set of magnetite samples was made by sizing particles obtained by mechanical pulverization of concentrates obtained from taconite ore. The sizing, into seven lots, was accomplished with the use of a Haultain infrasizer, the method being one of aerial elutriation. Microscope observation revealed the particles to be very angular in shape. Within each sample there was a fairly large size distribution. Table I gives the approximate mean diameter of the seven samples, as determined with an optical microscope. The percentage of Fe_3O_4 is a maximum value, since there may have been small amounts of other iron compounds present. The low values of magnetite in samples 1 to 3 were mainly due to silica impurities, owing to the nature of the ore and the concentrating process.

TABLE I.

Sample number	1	2	3	4	5	6	7
Mean diameter (microns)	50	44	31	21	14	4	1.5
%Fe_3O_4 (max)	73	86	89	96	97	96	93

Additional samples were made from pigment oxides. Yellow pigment oxide ($FeO \cdot OH$) is dehydrated to α-Fe_2O_3 and then reduced to Fe_3O_4 in a hydrogen atmosphere. The magnetite was also further oxydized to obtain γ-Fe_2O_3. The chemical process is similar to that used by Osmond.[6,7] The parent pigment particles are acicular in shape. Electron microscope photographs do not reveal any change in size or shape in any of the resulting products. An average particle is about 1μ long and has a length-to-width ratio of about 7 or 8. The lengths of the particles vary from about 0.1 μ to at most 2.0 μ. One electron microscope photograph of the pigment is shown in Fig. 5. It appears that some particles have a short needlelike extrusion occurring near the top of the acicular particle, and making an angle of about 60° with it. The percentage of particles that have such "thorns" is unknown. The acicular magnetite oxidize at room temperatures, so that the chemical analysis showed that the sample used had only 22% of ferrous iron instead of the stoichiometric 24.1%. The γ-Fe_2O_3 was 99% pure, the impurity probably being due mainly to H_2O.

On the basis of observations by Ervin on the transformation of α-$AlO \cdot OH$ to α-Al_2O_3, Osmond[7] infers that the corresponding acicular iron oxide particles probably consist of a collection of crystallites, with easy direction of magnetization parallel to the long dimension of the particle. Further, he concludes that each particle contains many small cavities between the crystallites, caused by loss of oxygen during the chemical changes. However, these cavities are so small that the particle still behaves as a unit magnetically.

Another pigment sample investigated was more of a cubical or equant shape. These particles were about 1 μ along each edge. The material employed in our experiments was about 97.5% pure, but had partially oxidized toward γ-Fe_2O_3, so that its ferrous iron content was 18.4%.

The calculation of the critical size predicts that all these powders except the acicular ones, have some sort of flux closure configuration, and are, in fact, likely to be multidomained. On the other hand, at least the majority of the acicular particles are predicted to be single domained.

FIG. 5. Electron microscope picture of typical acicular particles of $FeO \cdot OH$.

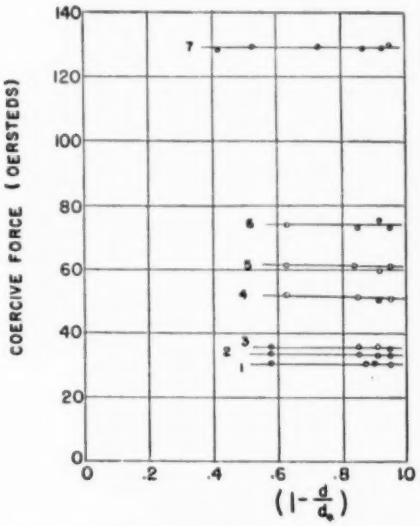


FIG. 6. The variation of coercive force with compression of magnetite powders obtained by pulverization and subsequent sizing.

As mentioned in the introduction, the measurements of coercive force were made as a function of the density of the powders. Compression of the taconite magnetite samples 1 to 6 caused breakage of the particles. Therefore, in order to observe the variation of coercive force with density, these samples were diluted with various amounts of barium sulfate. The samples were immobilized in paraffin, since otherwise rotation of some of the particles in the applied field gave values of the coercive force slightly too small. The powders consisting of smaller particles of Fe_3O_4 and γ-Fe_2O_3 were compressed in a brass die up to pressures of 60 000 lb psi.

The variation of coercive force as a function of $(1-d/d_0)$ for the various powders tested is shown in Figs. 6 and 7. Here d_0 is the density of the bulk material (5.2 g/cc for Fe_3O_4 and 5.0 g/cc for γ-Fe_2O_3), and d is the density of the sample. For the taconite magnetite sample 7, the first two points represent compressed material, the others represent material diluted with $BaSO_4$. It is clear that there is no significant variation of the coercive force with density of the taconite magnetite.

Further, the equant material did not show an appreciable decrease of H_c with compression, whereas the acicular Fe_3O_4 and γ-Fe_2O_3 showed a very significant decrease, (Fig. 7). It may be noted that there is a deviation of the experimental points at greatest powder compressions from the straight lines of best fit towards higher values of H_c. This is undoubtedly due either to the introduction of strains or else the breaking up of the particles.

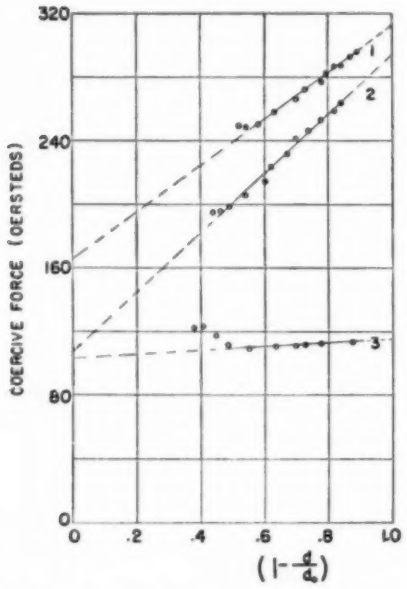


FIG. 7. The dependence of the coercive force on powder density for (1) acicular Fe_3O_4, (2) acicular γ-Fe_2O_3, and (3) cubical Fe_2O_4.

DISCUSSION

As pointed out by Néel,[11] a linear decrease of H_c with $(1-d/d_0)$ is expected for single domain particles whose shape is the factor governing the value of the coercive force. This decrease results because of the decrease of magnetostatic energy in the spaces between particles as these spaces get smaller after a compression. This result follows because for these single domain particles the coercive force is proportional to the energy. This result also holds for the case of holes in bulk material, according to the Néel-Kersten[12,13] theory.

On the other hand, the magnetostatic energy in the spaces between particles is much reduced if the particles are multidomained, since such domains are oriented to achieve the condition of minimum free energy. It is very possible that there are no free poles, and therefore no magnetostatic energy exists. In any event, it is certain that the domain configuration is such to reduce greatly the magnetostatic energy that would be produced if the particle consisted of a single domain. Consequently, it is to be expected that the coercive force for such powders would decrease but little, if at all, with compression. Hence, we conclude that the experimental results are consistent with the theoretical predictions of Fig. 4 that the majority of particles of the acicular powders are single domains, whereas all the other powders consist of multidomain particles.

We pointed out earlier that the value of the coercive force might be used as a criterion for the determination of the critical size. The coercive force of the acicular powders at infinite dilution $(1-d/d_0=1)$ is considerably less than that predicted by theory. Experimentally we observe 316 and 294 oersteds for the acicular Fe_3O_4 and γ-Fe_2O_3, respectively, as compared to the theoretical values of ~1170 and 976 oersteds derived, assuming the existence of only shape anisotropy. Too low a value for H_c for single domain particles has been found by a number of workers for various magnetic metal micropowders.[14] We will discuss possible reasons for these low values later.

The coercive force due only to crystalline anisotropy (145 oe for Fe_3O_4) is about the order of the coercive force observed experimentally for taconite sample No. 7 and for the equant powder sample. Further, the coercive force caused by crystalline anisotropy probably has a different dependence on density than the coercive force caused by shape anisotropy. In fact, it appears likely that there would be only a small, if any, dependence on packing. Therefore, use of the coercive force criterion would admit the possibility that the taconite sample No. 7 and the equant sample consist of single domain particles whose coercive force is almost entirely caused by crystalline anisotropy. However, since the particles of these powders are not perfect spheres, but have an appreciable shape factor, and because they are considerably larger than the critical size predicted theoretically, it is very likely they are multidomained. In general, the criterion based on particle size and shape is probably more reliable than the coercive force criterion for the determination of critical size.

There is another way a single domained powder may have a coercive force independent of packing. If the particle contains very large holes, or nonmagnetic inclusions, various regions of the single domain particle

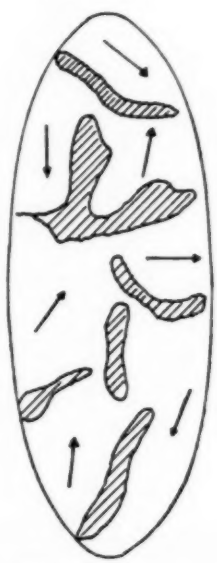


FIG. 8. Large cavities, in a single domain particle, that isolate regions so effectively that they behave independently magnetically.

[11] L. Néel, Compt. rend. **224**, 1550 (1947).
[12] L. Néel, Ann. univ. Grénoble **22**, 299 (1946).
[13] M. Kersten, *Grundlagen einer Theorie der Ferromagnetischen Hysterese und der Koerzitivkraft* (S. Hirzel, Leipzig; reprinted J. W. Edwards Brothers, Inc., Ann Arbor, 1943).
[14] W. H. Meiklejohn, Revs. Modern Phys. **25**, 302 (1953), for example.

may be separated in such a way that the particle acts as a multisingle domain particle. See Fig. 8.

The different slope of curves (1) and (2) of Fig. 7 results in two different, nonzero values of H_c at the extrapolation of $(1-d/d_0)$ to zero, that is, at the bulk density limit. These values probably represent a coercive force where the contribution due to shape has disappeared due to particle interaction. This is supported by the fact that the intercepts are close to the theoretical value of the coercive force due only to crystalline anisotropy (145 oe for Fe_3O_4).

Finally, we discuss qualitatively the reasons why the coercive force of the acicular powders is considerably less than that predicted by theory. While the shape anisotropy is a dominating factor with these powders, crystalline anisotropy and any strains that may exist cannot be neglected. The influence of these factors depends on the orientation of the easy directions of magnetization with respect to the long axis of the particle, and may increase or decrease the coercive force. Even if they are oriented in such a direction to produce the maximum reduction in H_c, they will not account for the large decrease observed. We have been led, therefore, to consider the internal cavities in the particles, and how they may reduce the coercive force. There appear to be two possibilities.

(1) The iron ions interact, in ferrites, via the oxygen ions in a process known as superexchange. If the holes are many and small, of the order of the radius of one oxygen ion, then there may be an appreciable decrease in the aligning force, and therefore a decrease in the saturation magnetization.

(2) The magnetostatic energy in the cavity may affect H_c by virtue of the fact that the internal magnetostatic energy is a function of the shape and orientation of the hole. If the polar axis of an ellipsoidal cavity lies almost along the polar axis of the particle, the coercive force will be increased. On the other hand, if the polar axis of the cavity lies approximately in the equatorial plane of the particle, H_c will be decreased.

CONCLUSIONS

The particle size below which magnetite or gamma ferric oxide particles exist with the single domain configuration has been calculated. For magnetite, the particles' major axis critical length for a prolate spheroid of 10/1 major to minor axis ratio is found to be about 3 μ. The critical sizes for γ-Fe_2O_3 are somewhat larger.

The variation of coercive force with the density of powders of variously shaped and sized iron oxide powders is consistent with the calculated single domain size. However, the absolute value of the coercive force of the single domain particles is much smaller than predicted by theory. Crystalline anisotropy and internal cavities may provide the explanation of this behavior.

ACKNOWLEDGMENTS

We are indebted to N. F. Schulz of the Mines Experimental Station for performing the infrasizing of the pulverized magnetite. The cubical and acicular powders were made available through the kind cooperation of W. W. Wetzel and A. H. Persoon of the Minnesota Mining and Manufacturing Company, St. Paul, Minnesota. The electron microscope photograph of the acicular pigment particles is included through the courtesy of C. K. Williams.

We appreciate very much the interest shown in this work by H. E. Hartig. Finally, we wish to thank E. W. Davis and H. E. Wade for their support of this investigation.

Letters to the Editor

Note on Electron Diffraction Patterns of CuO

R. D. Heidenreich and K. H. Storks
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey
(Received April 30, 1955)

In a recent paper, Cowley[1] has ascribed the intensity anomalies in electron diffraction patterns of CuO (first reported by Murison[2]) to a spike or needle-shaped crystal habit. These intensity anomalies are chiefly in the (110), $d=2.76$ A, and (112), $d=1.93$ A, reflections which are about four times and ten times stronger, respectively, than observed in x-ray diffraction patterns

(a) (b)

Fig. 1. Electron diffraction transmission pattern (a) and reflection pattern (b) from a CuO film formed on OFHC copper at about 400°C. 40 kv.

of Tenorite.[3] The patterns are commonly referred to as "3-ring copper oxide" or CuO′ and have been observed a number of times in these laboratories and are the rule for the oxide layer on pure copper oxidized in air at temperatures in the range 300–600°C. The CuO′ pattern has been obtained from both polycrystalline films and spikes or needles grown on copper surfaces. Figure 1 illustrates a transmission pattern obtained from an oxide film stripped from OFHC copper oxidized at 400°C and a reflection pattern from a similar film while still on the copper base. Table I lists the d values and visually estimated relative intensities for the

Fig. 2. Electron micrograph of the CuO film used for Fig. 1(a) illustrating absence of spikes or needle-shaped crystals.

Table I. d values and visually estimated relative intensities from Fig. 1(a).

hkl	d	Rel. intensity
(110)	2.76	4
(002), (1̄11)	2.51	8
(111), (200)	2.29	10
(1̄12)	1.93	3
(202)	1.82	4
	1.67	1
(202)	1.56	5
(113)	1.47	5

transmission pattern. The (110) is considerably weaker in the reflection pattern than by transmission while the (1̄12) is about the same. This may be due to some preferred orientation in the film. Figure 2 shows an electron micrograph of the stripped film and does not exhibit the spikes or needles as required by Cowley. It is concluded that crystal shape is not important in the intensity anomalies and that Cowley's explanation is not adequate. The origin of the anomalies must reside either in the structure itself or in the interaction of diffracted beams in the crystal (dynamical effects).

[1] J. M. Cowley, J. Electrochem. Soc. **101**, 277 (1954).
[2] C. A. Murison, Phil. Mag. **17**, 96 (1934).
[3] ASTM x-ray diffraction data card.

Sputtering of Metal Single Crystals by Ion Bombardment

Gottfried K. Wehner
Electronic Components Laboratory, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio
(Received February 28, 1955)

Experiments in the field of sputtering by ion bombardment yield reliable and basic information only under the following conditions: (1) mean free path of sputtered atoms larger than tube dimensions, (2) well-defined energy and angle of incidence of the bombarding ions, and (3) high density of the bombarding ion current or very clean discharge conditions such that sputtering can overcome the formation of impurity layers on the target surface.

Such conditions are achieved when the negative target is immersed in a low pressure supported mercury plasma of high density. With the target large compared to the ion sheath thickness, the incidence becomes essentially normal. Our studies were made in a demountable mercury pool cathode tube with the following data: Hg pressure 1 micron, ion current densities to the target of the order 10 ma/cm², target temperature ~200°C, and ion energies varied between threshold and several hundred ev. It was previously determined under similar conditions that the elastic constants of the target material (sound velocity) were influential on the sputtering threshold,[1] and it was shown that sputtering at low ion energies is basically a momentum transfer and not an evaporation process.[2] These results indicated that interesting observations could be expected from controlled sputtering of metal single crystals.

Material sputtered from tungsten single crystal balls (~1-mm diameter, 150 ev Hg ions) is deposited on a glass collector plate (1 cm distant) in patches which resemble electron field emission patterns (Fig. 1). A closer study reveals that the sputtered atoms are ejected independent of surface configuration or angle of incidence of the bombarding ions in the direction of close-packed atoms, hence [110] in Ag and Cu, [111] in W or α Fe. Material sputtered from a small area of a low index plane is therefore deposited in simple characteristic patterns. Figure 2 shows for Ag together with the corresponding crystal models in (a) the threefold symmetrical pattern obtained from a (111) plane, (b) a center patch surrounded by four patches in twofold symmetry as obtained from a (110) plane, and (c) the four patches from a

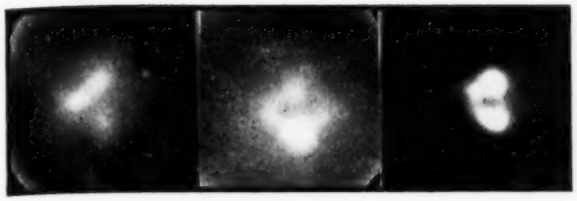


FIG. 1. Deposit patterns obtained from three different tungsten single-crystal balls when sputtered with 150 ev Hg ions.

(100) plane. Those directions which have the largest angle with the direction of incidence and therefore require the least directional change of momentum predominate. This checks with the former observation that atoms sputtered from polycrystalline material under oblique incidence of the bombarding ions are ejected in a direction away from the direction of incidence.[2]

Material which is sputtered parallel to the target surface is mostly trapped and causes growth of oriented hillocks. This effect is most pronounced in a f.c.c. (111) plane, where six close-packed directions are parallel to the surface and shown in Fig. 3, a micrograph of a (111) Ag plane with such characteristic hillocks developed after prolonged sputtering.

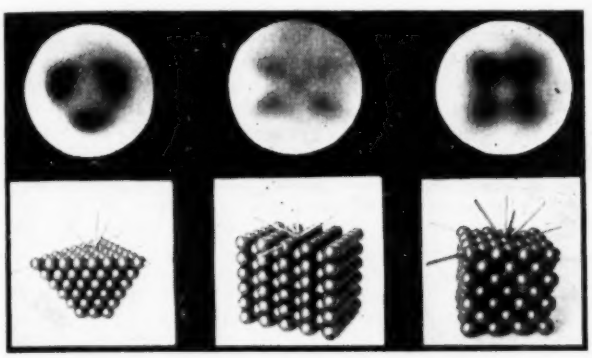


FIG. 2. Deposit patterns obtained from low index planes of a silver single crystal when sputtered with Hg ions. (a) left, (111) plane, 100 ev. (b) center, (110) plane, 50 ev. (c) right, (100) plane, 100 ev.

The changes which the deposited patterns undergo at different energies of the bombarding ions reveal interesting details of the sputtering process. For instance, near threshold only such atoms are sputtered which have a close-packed direction for ejection which is not obstructed by neighbor atoms. At higher ion energy additional atoms are freed from the (more numerous) positions where neighbor atoms may interfere with the direction of close-packed rows. This causes deviations from these directions and characteristic and predictable changes in these patterns.

The fact that the patterns are essentially independent of the surface configuration provides a simple and reliable method of studying surface lattice distortions and later, after the surface layers have been sputtered off, of roughly determining the orientation of a metal single crystal.

More details of these studies will be reported in the near future.

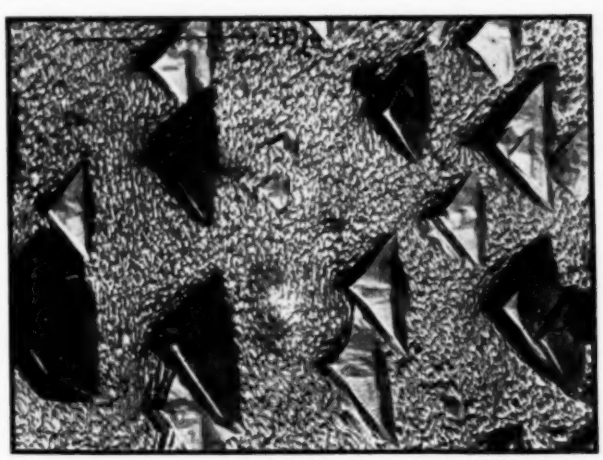


FIG. 3. Micrograph of (111) silver single-crystal plane after sputtering with Hg ions (125 ev) showing growth hillocks.

[1] G. Wehner, Phys. Rev. **93**, 633 (1954).
[2] G. Wehner, J. Appl. Phys. **25**, 270 (1954).

Interpretation of the Strain Broadening Components in X-Ray Diffraction Patterns

J. C. BOWMAN, J. A. KRUMHANSL, AND J. R. STOCK
National Carbon Research Laboratories, Cleveland, Ohio
(Received May 11, 1955)

WARREN and Averbach[1] have represented the broadened line shape of crystalline (00l) x-ray reflections from parallel layer type crystals as a Fourier series

$$P_{s\theta}=K\sum_n A_n(l)\cos 2\pi n h_3 \qquad (1)$$

with $h_3=(2a_3\sin\theta)/\lambda$. After removal of geometric and particle size[2] broadening effects from the coefficients $A_n(l)$, the remaining broadening is conveniently discussed in terms of $\langle\Delta L^2\rangle^{\frac{1}{2}}$, the root mean square distortion in length of a column of nominal length L. Experimentally, the quantity $\langle\Delta L^2\rangle^{\frac{1}{2}}/L$ is found to be a decreasing function of L, interpreted as a nonuniform "strain" decreasing for large L. The purpose of this note is to show that the observed effect is not really surprising but follows naturally from consideration of a model for random strains.

In introducing the argument, let us note two points: (a) Homogeneous strain is equivalent to a change in the lattice constant and appears as a line shift only. Thus shape analysis cannot distinguish uniform strain. (b) Purely random distortion necessarily broadens the line, and under certain circumstances also causes a shift. These general features appear from the following analysis.

A model for the strained crystal is fully specified by giving the displacements $Z_1(m)$ between adjacent parallel layers: $Z_1(m)=Z(m+1)-Z(m)$, where $Z(m)$ is the absolute position of the mth layer in units of the layer separation a_3. Consider the particular model in which

$$Z_1(m)=\epsilon+\delta_1(m), \qquad (2)$$

where ϵ is a constant giving the displacement representing uniform strain and $\delta_1(m)$ is the random component on m with average value zero. The $Z_1(m)$ are assumed statistically independent. For the displacement between two layers nominally n spacings apart we have

$$Z_n(m)=n\epsilon+\delta_n(m). \qquad (3)$$

For an arbitrary probability distribution for δ_1, the probability distribution for δ_n is given by n-fold multiple "Faltung." The Fourier coefficients of the reflected line are given by $A_n=\langle\cos 2\pi l Z_n\rangle$, $B_n=-\langle\sin 2\pi l Z_n\rangle$. For the present model these values, or rather the values of the Fourier magnitude $C_n=(A_n^2+B_n^2)^{\frac{1}{2}}$ and phase angle $\phi_n=\arctan(B_n/A_n)$, are calculated[3] to be

$$C_n=(C_1)^n \qquad (4)$$
$$\phi_n=-n(2\pi l\epsilon+\psi_1).$$

Here C_1 and ψ_1 are constants depending only upon the order of reflection l and the probability distribution for δ_1. The variation of ϕ_n with n shows the previously noted shift of the line with uniform strain. The C_n's determine the shape independently of position. Following the usual treatment of experimental line

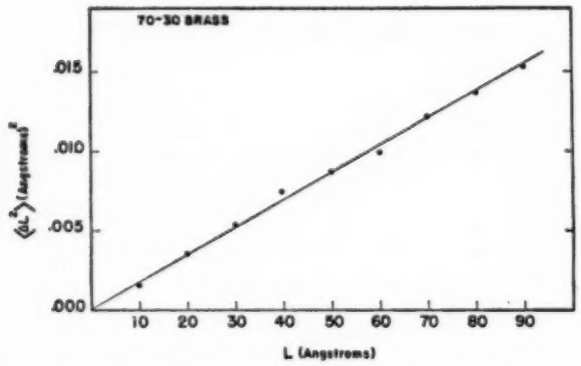


FIG. 1. Measured mean square displacement $\langle \Delta L^2 \rangle$ for columns of length L in a sample of cold worked 70–30 brass.

broadening data,

$$C_n \rightarrow \exp[-2\pi^2 l^2 \langle Z_n^2 \rangle_{\text{meas}}], \tag{5}$$

we find that $\langle Z_n^2 \rangle_{\text{meas}}$ is proportional to n and independent of ϵ. With $L=na_3$ and $\Delta L = Z_n a_3$, it is $\langle \Delta L^2 \rangle$, not $\langle \Delta L^2 \rangle^{\frac{1}{2}}$ which generally should be linear with L. Examination of data on α brass obtained at this laboratory (Fig. 1) and of other published data[4,5] shows satisfactory agreement with this model. This model also provides a more valid basis for the extrapolation of strain to small values of L, for example, in order to determine the strain energy. Further insight can be gained by comparing the true mean square value of Z_n, computed directly from Eq. (3), with the value obtained by way of the Fourier coefficients, Eq. (5):

$$\begin{aligned} \langle Z_n^2 \rangle_{\text{true}} &= n^2\epsilon^2 + n\langle \delta_1^2 \rangle, \\ \langle Z_n^2 \rangle_{\text{meas}} &= n\alpha \langle \delta_1^2 \rangle. \end{aligned} \tag{6}$$

Here α is a constant which is unity if the distribution for δ_1 is Gaussian or sufficiently narrow; otherwise α may deviate from one.

To summarize, the measured rms layer displacement gives no information about uniform strain; secondly, it is the mean square displacement which is proportional to column length L. Work in progress on several models with correlated $\delta_1(m)$ shows these same general features for large n but important deviations for small n.

[1] B. E. Warren and B. L. Averbach, J. Appl. Phys. **21**, 595 (1950).
[2] B. E. Warren and B. L. Averbach, J. Appl. Phys. **23**, 497 (1952).
[3] A similar analytical method has been used by B. E. Warren and C. R. Houska, J. Appl. Phys. **25**, 1503 (1954).
[4] B. E. Warren and B. L. Averbach, *Imperfections in Nearly Perfect Crystals* (John Wiley and Sons, Inc., New York, 1953).
[5] D. T. Keating, Phys. Rev. **97**, 832 (1955).

Thermal Effects in Point-Contact Rectifiers

R. E. BURGESS
The University of British Columbia, Vancouver, Canada
(Received April 6, 1955)

IN two recent letters[1,2] Armstrong purported to show that turnover in a point-contact rectifier could be explained on the basis of a reverse characteristic

$$i = a \exp(-c/k\theta),$$

where θ is the contact temperature which is greater than the ambient temperature T because of the heating effect of the applied power.

For a voltage extremum ($dV/dI=0$) we have the equation

$$\exp x = mx,$$

where $x=iv/2\pi$ gr and $m=c/kT \gg 1$. This equation has two roots x_1 and x_2 and the smaller root x_1 which is to a close approximation given by $1/(m-1)$ was chosen as the condition of turnover in Armstrong's first letter. Although this correctly corresponds to a voltage maximum, the predicted turnover power increases with increasing ambient temperature, in contradiction to the experimentally observed behavior.

In the second letter Armstrong suggested that the larger root x_2 was appropriate to the turnover condition since it gave a turnover power which decreased with increasing ambient temperature. However *this root is completely unacceptable since it corresponds to a voltage minimum.*

The writer[3] has recently shown that only rather special classes of function yield dependences of turnover current and voltage on ambient temperature which are consistent with experiment.

[1] H. L. Armstrong, J. Appl. Phys. **24**, 1332(L) (1953).
[2] H. L. Armstrong, J. Appl. Phys. **25**, 1345(L) (1954).
[3] R. E. Burgess, Proc. Phys. Soc. (London) (to be published).

Books Reviewed

Ionography. H. J. MCDONALD. Pp. 268+x. Year Book Publishers, Inc., Chicago, 1955. Price $6.50.

In reviewing the present knowledge regarding electromigration in stabilized media, the author and his collaborators have stressed the topic of paper-stabilized electrolytes. In addition, work on ion or particle migration in electrolytes stabilized with such materials as agar, silicic acid, gelatin, starch, ground-glass wool, glass beads, and cloth, cotton or silk, is treated. The subjects of paper chromatography and electrophoresis in nonstabilized electrolytes are not discussed except as they pertain to the main theme of the book. Since the experience of the authors centers mainly on the use of the horizontal paper strip in a gas atmosphere for conducting ionographic separations, this is the experimental technique which receives the greatest emphasis. In general, the graphs represent their own experimental work. A review of references to electrographic analysis and iontophoresis is provided.

Electronik des Einzelelektrons. FRANZ OLLENDORFF. Pp. 643+xii. Springer-Verlag, Vienna, 1955. Price $23.10.

This comprehensive work on electron ballistics is part of a series "Technische Electrodynamik" by Dr. Franz Ollendorff. The present book is the first part of volume II, "Innere Elektronik." The trajectories of electrons and ions in electric and magnetic fields are analyzed first from an elementary point of view and then with the generalized methods: Hamilton's principle, Larmor's theorem, canonical equations, Glaser's method, and the principle of least action. Electron optics, including aberration theory, is developed fully. The final section introduces relativistic electron ballistics and applies this to the planar magnetron, betatron, synchrocyclotron, synchrotron, and linear accelerator. Throughout the book many practical systems are treated, such as deflecting plates, multiplier phototubes, mass spectrographs, electron guns, and electron lenses.

Strength of Materials. Part I. Elementary Theory and Problems. S. TIMOSHENKO. Pp. 442+xiv. Third edition. D. Van Nostrand Company, Inc., New York, 1955. Price $6.00.

The third edition of this undergraduate engineering textbook provides many new problems and two new chapters. The new material includes bending in a plane which is not a plane of symmetry, the bending of curved bars, and the concept of shear center. The book is designed as a textbook for introductory courses in the strength of materials.

Optical Properties of Thin Solid Films. O. S. HEAVENS. Pp. 261+vii. Academic Press, Inc., New York, 1955. Price $6.80.

The formation, structure, optical properties, and applications of thin solid films are discussed in this book. The principal

applications treated are low-reflection coatings on glass, high-efficiency reflectors, interference filters, and thin-film polarizers. Techniques for measuring the thickness and optical constants of a film are described.

Cycles and Performance Estimation. Vol. I of Gas Turbine Series. JAMES HODGE. Pp. 329+xvi. Academic Press, Inc., New York, 1955. Price $9.00.

Although some discussion of other applications of gas turbine cycles is given, the principal part of this book is concerned with the theory of aircraft engine cycles. The book is intended for the design engineer.

Neutron Diffraction. G. E. BACON. Pp. 299+ix. Oxford University Press, New York, 1955. Price $5.60.

The entire field of neutron diffraction is discussed in this book, which is one of the series entitled "Monographs on the Physics and Chemistry of Materials." The aim of this monograph is to give a comprehensive account of the principles, practice, and achievements of neutron diffraction. It will probably be especially useful to students of the solid state since it presents many examples of the application of neutron diffraction to the study of the structures and magnetic properties of solids. About three-hundred references to the literature are provided.

Introduction to Physics. F. M. DURBIN. Pp. 780+xiv. Prentice-Hall, Inc., New York, 1955. Price $6.50.

The approach in this college text is from the liberal arts viewpoint. Intended for the standard eight-hour course, the book places less stress on practical applications than is usually found. Certain topics may be studied with or without a knowledge of trigonometry. Where such a mathematical treatment is necessary the reader is referred to the appendix. The first chapter is designed to familiarize the student with the application of formulas and provide practice in the analysis and solution of problems. The concepts of accelerated motion and Newton's laws are deferred till the eighth chapter. The first few chapters deal with discussions of scalar and vector quantities, equilibrium, work and energy, machines, mechanics of liquids and gases, and elasticity of solids, liquisd, and gases. In order to allow room for a treatment of modern physics some descriptive topics usually found in a physics text have been omitted and others presented as problems. The cgs system is used throughout.

Elements of Atomic Physics. D. C. PEASLEE. Pp. 394+ix. Prentice-Hall, Inc., New York, 1955. Price $7.50.

This book is intended as a text for an intermediate undergraduate course in atomic physics. The first seven chapters present the "facts of atomic structure," including geometrical, magnetic, optical, and electrical properties of atoms, and a description of optical and x-ray spectra. Chapters 8–12 provide the "theoretical interpretation of the facts" and discuss relativity, wave mechanics, and the application of wave mechanics to atomic structure. Chapters 13 and 14 present recent developments and quantum-statistical mechanics. There are about 140 problems for the student.

Fundamental Formulas of Physics. D. H. MENZEL. Pp. 765 +xxxv. Prentice-Hall Inc., New York, 1955. Price $10.65.

This book is intended as a reference book for the physicist who understands the methods of a particular branch of physics but who wishes to check or refresh his memory before using particular formulas. There is a relatively small amount of explanatory material in this volume, but in some sections the limitations of the mathematical expressions are carefully stated. Formulas from the fields of mathematics, physical chemistry, astrophysics, meteorology, and biophysics are included. The following sections will probably be of especial interest to the readers of this journal: Solid State (Herring), The Theory of Magnetism (Van Vleck), Electron Optics (Ramberg), and Electronics (Chaffee).

Elements of Physics for Students of Science and Engineering. G. SHORTLEY AND D. WILLIAMS. Pp. 898+xiv. Second edition. Prentice-Hall, Inc., New York, 1955. Price $10.60.

This text is designed for the student taking an introductory physics course concurrently with instruction in the calculus, the latter being introduced gradually as the student progresses. The book aims to provide ". . . a working knowledge of the fundamental principles that describe all physical phenomena, of how they evolved, and of their scope and limitations." The broad conservation principles are stressed.

Various revisions and additions have been made in this new edition. Subject matter in mechanics has been rearranged, illustrative examples included, and the problems have been improved, expanded, and placed at the ends of the chapters. New topics discussed include relative velocities, dynamical systems, fluid viscosity, surface phenomena in liquids, Gauss' law in electrostatics, thermionic vacuum tubes, transistors, and the particles of modern physics. There are six major divisions of the book; *Mechanics, Heat, Wave Motion and Sound, Light, Electricity and Magnetism,* and *Modern Physics.*

Mathematical Foundations of Quantum Mechanics. J. VON NEUMANN. Pp. 445+xii. Princeton University Press, Princeton, 1955. Price $6.00 (paper).

This book is a translation by R. T. Beyer of the classic work of von Neumann, originally published in German. It is well known that this book established a rigorous mathematical foundation under the new wave mechanics. The author has demonstrated that much of the early, wave mechanical theory was unnecessarily lacking in rigor and that the whole structure would be made sound by reliance on the Hilbert theory of operators.

Advanced Mathematics for Engineers. H. W. REDDICK AND F. H. MILLER. Pp. 548+xiv. Third edition. John Wiley and Sons, Inc., New York, 1955. Price $6.50.

Particular emphasis in this book has been placed upon the application of mathematical principles to engineering and physics problems. Each chapter begins with sections describing the mathematical principles and concludes with a large number of applications. Rigor is not emphasized. A large number of problems to be solved (with answers) is found at the end of each chapter. This volume contains treatments of ordinary and partial differential equations, infinite series, Fourier series, gamma, Bessel, and Legendre functions, vector analysis, and complex variables. Chapters on elliptic and hyperbolic functions, probability theory, and operational methods as applied to engineering are included. The appendix describes dimensional analysis and the physical units used throughout the text. In the third edition the problem sets have been expanded, and a section on LaPlace transforms and a discussion of numerical methods for solving differential equations have been inserted.

Electrochemistry in Biology and Medicine. T. SHEDLOVSKY, editor. Pp. 369+xii. John Wiley and Sons, Inc., New York, 1955. Price $10.50.

This book originated from the Symposium on Electrochemistry in Biology and Medicine held in New York during the meeting of the Electrochemical Society in April, 1953. Some of the nineteen chapters are essentially like the papers presented, and others are the result of further deliberation. The book contains the views of twenty-three experts and represents the trend of modern thinking in this field. Among

the contributors are physicists, chemists, biologists, and medical men who share a common interest in the electrochemical problems of living processes. The topics discussed include membranes, nerve and plant cells, biologically important ions, and applications of polarography, electrocardiography, and electroencephalography in medicine.

Books Received

Coordination, Control, and Financing of Industrial Research. ALBERT H. RUBENSTEIN, Editor. Pp. 429+xv. King's Crown Press, Columbia University, New York, 1955.

Technische Electrodynamik. Band II. Innere Elektronik, Teil 1: Elektronik des Einzelelektrons. FRANZ OLLENDORFF. Pp. 643+xii, Figs. 313. Springer-Verlag, Wien, 1955. Price $23.10.

General Physics. O. BLACKWOOD AND WILLIAM KELLY. Pp. 704+x, Figs. John Wiley and Sons, Inc., New York, 1955. Second edition. Price $6.75.

Petrographic Mineralogy. ERNEST E. WAHLSTROM. Pp. 408+vii, Figs. John Wiley and Sons, Inc., New York, 1955. Price $7.75.

General Principles of Geology. J. F. KIRKALDY. Pp. 327, Figs. Philosophical Library, Inc., New York, 1955. Price $6.00.

Announcements

Gaseous Electronics Conference

The Eighth Annual Gaseous Electronics Conference will be held at the General Electric Research Laboratory, The Knolls, Schenectady, New York, on Thursday, Friday, and Saturday, October 20, 21, and 22. For further information please write to Dr. J. D. Cobine, Secretary, Gaseous Electronics Conference Executive Committee, General Electric Research Laboratory, Schenectady, New York.